

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2002, 00:35:50 ; Search time 1937.48 Seconds
(without alignments)
5426.703 Million cell updates/sec

Title: US-09-809-524b-1

Perfect score: 779
Sequence: 1 atggcacacaggtaataaga.....gactggtttccatctgtatg 779

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estin:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_estc:.*
9: gb_est1:.*
10: gb_est2:.*
11: gb_hic:.*
12: gb_gss:.*
13: em_gss_hum:.*
14: em_gss_inv:.*
15: em_gss_pln:.*
16: em_gss_vrt:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	47.4	6.1	788	12	CNS00AK8
c 2	47.2	6.1	903	12	CNS00IIC
c 3	46.2	5.9	240	9	AU071288
c 4	46.2	5.9	240	9	AU072338
c 5	46.2	5.9	653	10	C24613
c 6	45.2	5.8	878	12	CNS0187R
c 7	44.4	5.7	568	12	CNS010PL
c 8	43.8	5.6	1101	12	CNS00KK2
c 9	43.6	5.6	1053	12	CNS016CV
c 10	41.8	5.4	908	12	CNS04KS8
c 11	41.8	5.4	939	12	CNS00CNG
c 12	41.8	5.4	1300	10	BM468018
c 13	41.4	5.3	536	12	AZ385473
c 14	41.2	5.3	888	12	CNS027EX
c 15	40.8	5.2	421	9	AV865711
c 16	40.8	5.2	453	9	AV865286
c 17	40.5	5.2	711	12	CNS011V8

c 18	40.6	5.2	842	12	CNS04993
c 19	40.6	5.2	1100	12	CNS012T7
c 20	40.4	5.2	511	12	TA116D02Q
c 21	40.4	5.2	996	12	CNS00FUH
c 22	40.2	5.2	439	12	CNS017HP
c 23	39.8	5.1	694	12	BH502270
c 24	39.8	5.1	969	12	CNS00IDL
c 25	39.6	5.1	264	9	AV888582
c 26	39.6	5.1	1101	12	CNS001FK
c 27	39.4	5.1	604	9	BE224195
c 28	39.4	5.1	1101	12	CNS00LOO
c 29	39.4	5.1	1147	12	B13042
c 30	39.2	5.0	668	9	AL514901
c 31	39.2	5.0	823	12	CNS0105K
c 32	39.2	5.0	1101	12	CNS00R39
c 33	39	5.0	517	10	BF193908
c 34	39	5.0	518	10	BJ079195
c 35	39	5.0	563	12	AQ319368
c 36	38.8	5.0	721	12	AZ197220
c 37	38.8	5.0	859	12	CNS004YY
c 38	38.8	5.0	1007	12	CNS06X9S
c 39	38.8	5.0	1084	12	CNS0025S
c 40	38.8	5.0	1101	12	CNS00L6N
c 41	38.6	5.0	463	9	AV595306
c 42	38.6	5.0	549	12	TA114H11Q
c 43	38.6	5.0	769	12	CNS04POW
c 44	38.6	5.0	838	12	CNS0096Q
c 45	38.6	5.0	993	12	CNS005N6

ALIGNMENTS

RESULT 1
CNS00AK8/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR21G10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL055582.1 GI:4936452
VERSION
KEYWORDS GSS.
SOURCE Drosophila melanogaster
ORGANISM fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Keizutyo Osoegawa and Aaron Mammoe in Pleter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BCPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1. 788
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"


```
Query Match      5.9%; Score 46.2; DB 9; Length 240;
Best Local Similarity 55.2%; Pred. No. 2.7;
Matches 90; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 5 cacacaggtaaatgagtgcttattgcattggctggtgcttttttaccgacattaaagcc 64
Db 59 CACAATTAACAAAGAGATGCTTGATGCAATGGATGATGCTGGAACCTGGAACATCAAGAC 118
Qy 65 ttctttattactccctattataaaagtgcagtttagggcgctcttttttaagtaatg 124
Db 119 TTGCTCTAATGCTTAAATGATTAATAAAGTTAGTATCGGTACTGATATTTCAAAATATTT 178
Qy 125 gcgtcttatgtcacttacccttcccatattaccacatttta 167
Db 179 CAATGCGCTGGTCTTTTATCTTGCCTGCAAAATCAACTCTAAGTTA 221

RESULT 4
AU072338 240 bp mRNA linear EST 23-JUN-1999
LOCUS AU072338 Dictyostelium discoidium SS (H.Urushihara) Dictyostelium
DEFINITION discoidium cDNA clone SSE223, mRNA sequence.
ACCESSION AU072338.1 GI:5162527
VERSION EST.
KEYWORDS Dictyostelium discoidium.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS Urushihara, H.
TITLE Developmental cDNA in Dictyostelium discoidium (1999)
JOURNAL Unpublished (1999)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = Dictyostelium discoidium cDNA project in Japan.
FEATURES
source
Location/Qualifiers
1..240
/organism="Dictyostelium discoidium"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSE223"
/clone_lib="Dictyostelium discoidium SS (H.Urushihara)"
/dev_stage="slug"

BASE COUNT 90 a 36 c 42 g 72 t
ORIGIN

Query Match      5.9%; Score 46.2; DB 9; Length 240;
Best Local Similarity 55.2%; Pred. No. 2.7;
Matches 90; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 5 cacacaggtaaatgagtgcttattgcattggctggtgcttttttaccgacattaaagcc 64
Db 66 CACAATTAACAAAGAGATGCTTGATGCAATGGATGATGCTGGAACCTGGAACATCAAGAC 125
Qy 65 ttctttattactccctattataaaagtgcagtttagggcgctcttttttaagtaatg 124
Db 126 TTGCTCTAATGCTTAAATGATTAATAAAGTTAGTATCGGTACTGATATTTCAAAATATTT 185
Qy 125 gcgtcttatgtcacttacccttcccatattaccacatttta 167
Db 186 CAATGCGCTGGTCTTTTATCTTGCCTGCAAAATCAACTCTAAGTTA 228

RESULT 5
C24613 653 bp mRNA linear EST 28-APR-1999
LOCUS C24613 Dictyostelium discoidium SS (H.Urushihara) Dictyostelium
DEFINITION discoidium cDNA clone SL-X008, mRNA sequence.
```

```
ACCESSION C24613
VERSION C24613.1 GI:2243034
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mitra, B. N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochial, H. and Tanaka, Y.
TITLE The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoidium cDNA project in Japan'
POLYA-No.

FEATURES
source
Location/Qualifiers
1..653
/organism="Dictyostelium discoidium"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SL-X008"
/clone_lib="Dictyostelium discoidium SS (H.Urushihara)"
/dev_stage="slug"

BASE COUNT 262 a 83 c 105 g 203 t
ORIGIN

Query Match      5.9%; Score 46.2; DB 10; Length 653;
Best Local Similarity 55.2%; Pred. No. 2.7;
Matches 90; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 5 cacacaggtaaatgagtgcttattgcattggctggtgcttttttaccgacattaaagcc 64
Db 127 CACAATTAACAAAGAGATGCTTGATGCAATGGATGATGCTGGAACCTGGAACATCAAGAC 186
Qy 65 ttctttattactccctattataaaagtgcagtttagggcgctcttttttaagtaatg 124
Db 187 TTGCTCTAATGCTTAAATGATTAATAAAGTTAGTATCGGTACTGATATTTCAAAATATTT 246
Qy 125 gcgtcttatgtcacttacccttcccatattaccacatttta 167
Db 247 CAATGCGCTGGTCTTTTATCTTGCCTGCAAAATCAACTCTAAGTTA 289

RESULT 6
CNS0187R/c 878 bp DNA linear GSS 26-JUL-1999
LOCUS CNS0187R Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN04E04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL108993
VERSION AL108993.1 GI:5629297
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
TITLE 1 (bases 1 to 878)
JOURNAL Genoscope.
COMMENT Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
```


Tetraodontidae; Tetraodon.
1 (bases 1 to 888)
Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 888)
Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 888)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1..888
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="242D06"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG242DB03Spl-end :
PUC-ori"
BASE COUNT 64 a 64 c 104 g 586 t 70 others
ORIGIN
Query Match 5.3%; Score 41.2; DB 12; Length 888;
Best Local Similarity 40.1%; Pred. No. 35;
Matches 157; Conservative 19; Mismatches 216; Indels 0; Gaps 0;
QY 26 ttatgcatggtgctggcgtttattcgaccattgaagcctttcttatttacttccctat 85
Db 361 ttt 420
QY 86 taaaagtgccagtttagggcgctcttttacgtaagcgttcttcttcttcttcttctt 145
Db 421 ttt 480
QY 146 tcccatattaccatcattaccagcagaagattatgatgcataattggtaaagattaca 205
Db 481 tntntntntnwtgattgttttttttttttttttttttttttttttttttttttttt 540
QY 206 gtggttaggttagtcaccgagaggtgattattggttttttttttttttttttttt 265
Db 541 ttttttttgcgataatgatkgtgagtggttggawkttttttttttttttttttttt 600
QY 266 cggttcccttttggcgttgatagggcgtttctctgtgacttcttcttcttcttctt 325
Db 601 ttttttttttttgggttttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
QY 326 caatgggtacgatattcattcacaatagaagcctcaactcttttggcttgccttt 385
Db 661 ttttttttttttggatatt 720
QY 386 tcagccagtttttgggtgttttttttttttttttttttttttttttttttttttttt 417
Db 721 ttt 752

RESULT 15
AV865711
LOCUS
DEFINITION AV865711 Nori Satoh unpublished cDNA library, egg Ciona
intestinalis cDNA clone rieg43f23 3', mRNA sequence.

AV865711 GI:16853235
EST.
Ciona intestinalis.
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
1 (bases 1 to 421)
Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
Expressed genes in Ciona intestinalis
Unpublished genes in Ciona intestinalis
JOURNAL
COMMENT
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..421
/organism="Ciona intestinalis"
/db_xref="taxon:7719"
/clone="rcieg43f23"
/clone_lib="Nori Satoh unpublished cDNA library, egg"
/tissue_type="whole animal"
/dev_stage="egg"
BASE COUNT 95 a 83 c 80 g 162 t 1 others
ORIGIN
Query Match 5.2%; Score 40.8; DB 9; Length 421;
Best Local Similarity 51.1%; Pred. No. 42;
Matches 96; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 304 cttagatacttttacgtggcgacaaatggggtacgatattcaattctacaatagaagctgaa 363
Db 221 CATGATATCAGTCAGTGGTCTCCAGCATGGCTATAATCTCTAATGCAACGCTTGTATCCGAA 280
QY 364 acctcaccttttggcgttcttccagccagtttttgggtgttttttttttttttttttt 423
Db 281 ATAGTCCCGTCTTGTGATTATTGCAAAAGTCGATGTGAAGTTGAACATACAGTCGGA 340
QY 424 ggcattgggtttattataaacaattctgtatagtcataatcaattttaccaccaggcgct 483
Db 341 GGCATTTCTGTTTCTCACATAAATACATATTAACAATTTGTTTAGATACCATCATTCAGA 400
QY 484 actttatt 491
Db 401 ATTTAAAT 408

Search completed: August 2, 2002, 02:16:16
Job time: 6026 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2002, 01:43:55 ; Search time 3427.07 Seconds
(without alignments)
4917.783 Million cell updates/sec

Title: US-09-809-524B-1
Perfect score: 779
Sequence: 1 atggcacacacggtaataaga.....gactggttcacatctgtatg 779

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 1081749327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_Main.*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US100_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US101_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	779	100.0	779	31	US-09-809-524A-1
2	779	100.0	779	31	US-09-809-524A-1
3	779	100.0	779	31	US-09-809-524B-1
4	763	97.9	779	31	US-09-809-524A-2
5	763	97.9	779	31	US-09-809-524A-2
6	763	97.9	779	31	US-09-809-524B-2
7	750.6	96.4	13417	10	US-08-637-759A-37
8	750.6	96.4	13417	10	US-08-637-759A-37
9	750.6	96.4	13417	12	US-08-637-759A-37
10	207.4	26.6	297	10	US-08-637-759A-15
11	207.4	26.6	297	10	US-08-637-759A-15
12	207.4	26.6	297	12	US-08-637-759A-15
13	118	15.1	43360	18	US-09-453-702B-206
14	118	15.1	43360	18	US-09-453-702B-261
15	76	9.8	70559	18	US-09-409-800B-1
16	63.6	8.2	810	21	US-09-543-681A-2303
17	60.2	7.7	801	33	US-09-868-604-33
18	59.4	7.6	1086	16	US-09-252-991A-13855
19	59.4	7.6	1365	16	US-09-252-991A-13593
20	59.4	7.6	1740	16	US-09-252-991A-13753
21	59.2	7.6	24153	47	US-60-082-302-682
22	53.6	6.9	16950	18	US-09-453-702B-166
23	53	6.8	771	36	US-09-974-300-1867
24	52.2	6.7	184	10	US-08-637-759A-16
25	52.2	6.7	184	10	US-08-637-759A-16
26	52.2	6.7	184	12	US-08-637-759A-16
27	48	6.2	2000	33	US-09-887-272A-5263
28	47.2	6.1	7218	8	US-08-466-194-14
29	43.4	5.6	603	17	US-09-397-761A-2950
30	43.4	5.6	603	32	US-09-838-601-2950
31	41.4	5.3	625	27	US-09-680-598A-688

Db 361 gaaacccacatttggctgctttcagccagtttttgctgtattttctttataagc 420
Qy 421 gggcgatgaggtttatataaacattctgtatgagtcataataatttaccaccagg 480
Db 421 gggcgatgaggtttatataaacattctgtatgagtcataataatttaccaccagg 480
Qy 481 cgtactttattttgaccggcaatttttaaaatatataccaggcagagtggaagacgctt 540
Db 481 cgtactttattttgaccggcaatttttaaaatatataccaggcagagtggaagacgctt 540
Qy 541 tatcaaatatgtcagttctctctccctgcccataataatgattgatttagccgactg 600
Db 541 tatcaaatatgtcagttctctctccctgcccataataatgattgatttagccgactg 600
Qy 601 gcttaggtcttttaaatcggtcgccacaaattggaattgtttttctctcaatgcg 660
Db 601 gcttaggtcttttaaatcggtcgccacaaattggaattgtttttctctcaatgcg 660
Qy 661 ctcaaaagtattggttctactgacgctcctgatctcattcccttatgctcttcacac 720
Db 661 ctcaaaagtattggttctactgacgctcctgatctcattcccttatgctcttcacac 720
Qy 721 tatttggtgaaagcgataattttatttattctataaagactggtttccatctgtatg 779
Db 721 tatttggtgaaagcgataattttatttattctataaagactggtttccatctgtatg 779

RESULT 3

US-09-809-524B-1

; Sequence 1, Application US/09809524B

; GENERAL INFORMATION:

; APPLICANT: Lowery, David

; APPLICANT: Kennedy, Michael J

; TITLE OF INVENTION: Salmonella Vaccine Materials and Methods

; FILE REFERENCE: 28341/6114.N

; CURRENT APPLICATION NUMBER: US/09/809,524B

; CURRENT FILING DATE: 2001-09-17

; PRIOR APPLICATION NUMBER: 60/190,178

; PRIOR FILING DATE: 2000-03-17

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 779

; TYPE: DNA

; ORGANISM: Salmonella dublin

US-09-809-524B-1

Query Match 100.0%; Score 779; DB 31; Length 779;
Best Local Similarity 100.0%; Pred. No. 3.3e-186;
Matches 779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggcacaacaggtaaatgagtgcttattgcatgctgctgtggtttttattcgaccatta 60
Db 1 atggcacaacaggtaaatgagtgcttattgcatgctgctgtggtttttattcgaccatta 60
Qy 61 agcctttctttattacttcccttattataaaggcgagtggttagggcgctcttttaagt 120
Db 61 agcctttctttattacttcccttattataaaggcgagtggttagggcgctcttttaagt 120
Qy 121 aatggcgctgtattgctacattacattccattatatacaatcattaccaggcagaagatt 180
Db 121 aatggcgctgtattgctacattacattccattatatacaatcattaccaggcagaagatt 180
Qy 181 atgctgcatattggttaagattacagttgttaggttaggttaggttaggttaggttaggt 240
Db 181 atgctgcatattggttaagattacagttgttaggttaggttaggttaggttaggttaggt 240
Qy 241 ggttttttaattgggtttttgtggtgggtttcccttttggggcggttgataggggggttt 300
Db 241 ggttttttaattgggtttttgtggtgggtttcccttttggggcggttgataggggggttt 300

Qy 301 ctgcttgataacttttacgtggtcgacaaatgggtacgatattcaattctacaaatagaagct 360
Db 301 ctgcttgataacttttacgtggtcgacaaatgggtacgatattcaattctacaaatagaagct 360
Qy 361 gaaacccacatttggctgctttcagccagtttttgctgtattttctttataagc 420
Db 361 gaaacccacatttggctgctttcagccagtttttgctgtattttctttataagc 420
Qy 421 gggcgatgaggtttatataaacattctgtatgagtcataataatttaccaccagg 480
Db 421 gggcgatgaggtttatataaacattctgtatgagtcataataatttaccaccagg 480
Qy 481 cgtactttattttgaccggcaatttttaaaatatataccaggcagagtggaagacgctt 540
Db 481 cgtactttattttgaccggcaatttttaaaatatataccaggcagagtggaagacgctt 540
Qy 541 tatcaaatatgtcagttctctctccctgcccataataatgattgatttagccgactg 600
Db 541 tatcaaatatgtcagttctctctccctgcccataataatgattgatttagccgactg 600
Qy 601 gcttaggtcttttaaatcggtcgccacaaattggaattgtttttctctcaatgcg 660
Db 601 gcttaggtcttttaaatcggtcgccacaaattggaattgtttttctctcaatgcg 660
Qy 661 ctcaaaagtattggttctactgacgctcctgatctcattcccttatgctcttcacac 720
Db 661 ctcaaaagtattggttctactgacgctcctgatctcattcccttatgctcttcacac 720
Qy 721 tatttggtgaaagcgataattttatttattctataaagactggtttccatctgtatg 779
Db 721 tatttggtgaaagcgataattttatttattctataaagactggtttccatctgtatg 779

RESULT 4

US-09-809-524-2

; Sequence 2, Application US/09809524

; GENERAL INFORMATION:

; APPLICANT: Pharmacia & Upjohn Company

; TITLE OF INVENTION: Salmonella Vaccine Materials and Methods

; FILE REFERENCE: 28341/6114.N

; CURRENT APPLICATION NUMBER: US/09/809,524

; CURRENT FILING DATE: 2001-03-15

; PRIOR APPLICATION NUMBER: 60/190,178

; PRIOR FILING DATE: 2000-03-17

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 779

; TYPE: DNA

; ORGANISM: Salmonella typhimurium

US-09-809-524-2

Query Match 97.9%; Score 763; DB 31; Length 779;
Best Local Similarity 98.7%; Pred. No. 3.7e-182;
Matches 769; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 atggcacaacaggtaaatgagtgcttattgcatgctgctgtggtttttattcgaccatta 60
Db 1 atggcacaacaggtaaatgagtgcttattgcatgctgctgtggtttttattcgaccatta 60
Qy 61 agcctttctttattacttcccttattataaaggcgagtggttagggcgctcttttaagt 120
Db 61 agcctttctttattacttcccttattataaaggcgagtggttagggcgctcttttaagt 120
Qy 121 aatggcgctgtattgctacattacattccattatatacaatcattaccaggcagaagatt 180
Db 121 aatggcgctgtattgctacattacattccattatatacaatcattaccaggcagaagatt 180
Qy 181 atgctgcatattggttaagattacagttgttaggttaggttaggttaggttaggttaggt 240
Db 181 atgctgcatattggttaagattacagttgttaggttaggttaggttaggttaggttaggt 240

```
QY 241 ggttttttaattggttttggcggttcccttttggcggttgcgttgatagcggtgttt 300
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 241 ggtttttcaattggttttggcggttcccttttggcggttgcgttgatagcggtgttt 300
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 301 ctgcttgatactttacgttggcgacaaatgggtacgatatcaattctacaataagaact 360
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 301 ctgcttgatactttacgttggcgacaaatgggtacgatatcaattctacaataagaact 360
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 361 gaaacctacttttggcttggcttttcagccagttttgtgtgtatttttcttttaaac 420
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 361 gaaacctacttttggcttggcttttcagccagttttgtgtgtatttttcttttaaac 420
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 421 ggcggcatgagtttatataaacattctgtatgagtcatacaatattaccaccagg 480
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 421 ggcggcatgagtttatataaacattctgtatgagtcatacaatattaccaccagg 480
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 481 cgtactttattgtcagtttctctctctgcataataatgtaggtattagccgatctg 600
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 481 cgtactttattgtcagtttctctctctgcataataatgtaggtattagccgatctg 600
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 541 tatcaattatgtcagtttctctctctgcataataatgtaggtattagccgatctg 660
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 541 tatcaattatgtcagtttctctctctgcataataatgtaggtattagccgatctg 660
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 601 gtttttagtcttttaaatcggtcgacacaaatggaatgtttttcttcaatgcg 660
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 601 gtttttagtcttttaaatcggtcgacacaaatggaatgtttttcttcaatgcg 660
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 661 ctcaaaagtatatgtgttctactgacgtctctgacatctcattcccttcttcaac 720
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 661 ctcaaaagtatatgtgttctactgacgtctctgacatctcattcccttcttcaac 720
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 721 tatttggtgaaagcgataaaatttatatttatctcaaaagactggttccatctgatg 779
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 721 tatttggtgaaagcgataaaatttatatttatctcaaaagactggttccatctgatg 779
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||

RESULT 5
US-09-809-524A-2
; Sequence 2, Application US/09809524A
; GENERAL INFORMATION:
; APPLICANT: Lowery, David
; TITLE OF INVENTION: Salmonella Vaccine Materials and Methods
; FILE REFERENCE: 28341/6114.N
; CURRENT APPLICATION NUMBER: US/09/809,524A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/190,178
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-09-809-524A-2
```

Query Match 97.9%; Score 763; DB 31; Length 779;
Best Local Similarity 98.7%; Pred. No. 3.7e-182;
Matches 769; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
QY 1 atggcacacacaggttaaatgagtgcttattgcattggctgtggttttttcgaccatta 60
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 1 atggcacacacaggttaaatgagtgcttattgcattggctgtggttttttcgaccattg 60
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 61 agcctttcttttactctccctattaaaagtggtgagtttagggcggtctcttttaagt 120
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 61 agcctttcttttactctccctattaaaagtggtgagtttagggcggtctcttttaagt 120
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 121 aatggcggtctatgctacattcttccctattaccatcattaccacagcaagaatt 180
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 121 aatggcggtctatgctacattcttccctattaccatcattaccacagcaagaatt 180
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
```

```
QY 181 atgatgcataattggttaaaagattacagttggttaggttagtcacggagaggtgattatt 240
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 181 atgatgcataattggttaaaagattacagttggttaggttagtcacggagaggtgattatt 240
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 241 gtttttttaattggtttttgtggttcccttttggcggttgcgttgatagcggtgttt 300
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 241 gttttttcaattggtttttgtggttcccttttggcggttgcgttgatagcggtgttt 300
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 301 ctgcttgatactttacgttggcgacaaatgggtacgatatcaattctacaataagaact 360
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 301 ctgcttgatactttacgttggcgacaaatgggtacgatatcaattctacaataagaact 360
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 361 gaaacctacttttggcttggcttttcagccagttttgtgtgtatttttcttttaaac 420
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 361 gaaacctacttttggcttggcttttcagccagttttgtgtgtatttttcttttaaac 420
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 421 ggcggcatgagtttatataaacattctgtatgagtcatacaatattaccaccagg 480
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 421 ggcggcatgagtttatataaacattctgtatgagtcatacaatattaccaccagg 480
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 481 cgtactttattgtcagtcgacaaatttttaaaatataccagcgagatggagaacgctt 540
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 481 cgtactttattgtcagtcgacaaatttttaaaatataccagcgagatggagaacgctt 540
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 541 tatcaattatgtcagtttctctctctgcataataatgtaggtattagccgatctg 600
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 541 tatcaattatgtcagtttctctctctgcataataatgtaggtattagccgatctg 600
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 601 gtttttagtcttttaaatcggtcgacacaaatggaatgtttttcttcaatgcg 660
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 601 gtttttagtcttttaaatcggtcgacacaaatggaatgtttttcttcaatgcg 660
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 661 ctcaaaagtatatgtgttctactgacgtctctgacatctcattcccttcttcaac 720
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 661 ctcaaaagtatatgtgttctactgacgtctctgacatctcattcccttcttcaac 720
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 721 tatttggtgaaagcgataaaatttatatttatctcaaaagactggttccatctgatg 779
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 721 tatttggtgaaagcgataaaatttatatttatctcaaaagactggttccatctgatg 779
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
```

RESULT 6
US-09-809-524B-2
; Sequence 2, Application US/09809524B
; GENERAL INFORMATION:
; APPLICANT: Lowery, David
; TITLE OF INVENTION: Salmonella Vaccine Materials and Methods
; FILE REFERENCE: 28341/6114.N
; CURRENT APPLICATION NUMBER: US/09/809,524B
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/190,178
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-09-809-524B-2

Query Match 97.9%; Score 763; DB 31; Length 779;
Best Local Similarity 98.7%; Pred. No. 3.7e-182;
Matches 769; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
QY 1 atggcacacacaggttaaatgagtgcttattgcattggctgtggttttttcgaccatta 60
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
Db 1 atggcacacacaggttaaatgagtgcttattgcattggctgtggttttttcgaccattg 60
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
QY 61 agcctttcttttactctccctattaaaagtggtgagtttagggcggtctcttttaagt 120
||||| |||||| |||||| |||||| |||||| |||||| |||||| |||||| ||||||
```

```
Db 61 agccttttcttacttcccttataaaagtgccagtttagggccgcacttttaagt 120
Qy 121 aatggcggttatgtcacttaccctttcccatattaccatcattaccagcagaagatt 180
Db 121 aatggcggttatgtcacttaccctttcccatattaccatcattaccagcagaagatt 180
Qy 181 atgatgatattggttaagattacagttggttagggtagtcaacggagagaggtattatt 240
Db 181 atgatgatattggttaagattacagttggttagggtagtcaacggagagaggtattatt 240
Qy 241 ggtttttaaattgggttttgccggcggttcccttttggcggttgatagtcgggggttt 300
Db 241 ggttttcaaattgggttttgccggcggttcccttttggcggttgatagtcgggggttt 300
Qy 301 ctgcttgatactttacgtggcgcgacaatgggtacgattatccaattctacaatagaagct 360
Db 301 ctgcttgatactttacgtggcgcgacaatgggtacgattatccaattctacaatagaagct 360
Qy 361 gaaacctcaacttttggcgttcttccagccagtttttgggtgtgtatttttttataagc 420
Db 361 gaaacctcaacttttggcgttcttccagccagtttttgggtgtgtatttttttataagc 420
Qy 421 ggcggcgatgggtttatattaaacattctgtatgagtcataatcaattttaccaccagg 480
Db 421 ggcggcgatgggtttatattaaacattctgtatgagtcataatcaattttaccaccagg 480
Qy 481 cgtactttattattgacccgcaattttttaaataatatccaggcagagtggaacgctt 540
Db 481 cgtactttattattgacccgcaattttttaaataatatccaggcagagtggaacgctt 540
Qy 541 tatcaattatgtcagtttctctctctccagccataatataatgattagtcagcagatctg 600
Db 541 tatcaattatgtcagtttctctctctccagccataatataatgattagtcagcagatctg 600
Qy 601 gcttaggtctcttaaaacggcgccacaaatgaatggtgtttttctctccaatgcg 660
Db 601 gcttaggtctcttaaaacggcgccacaaatgaatggtgtttttctctccaatgcg 660
Qy 661 ctcaaaagatatgggttctactgacgctcctgatctcattcccttatgctcttcacac 720
Db 661 ctcaaaagatatgggttctactgacgctcctgatctcattcccttatgctcttcacac 720
Qy 721 tatttggtgaagcgataaattttattttatttataaaagactggttccatctgtatg 779
Db 721 tatttggtgaagcgataaattttattttatttataaaagactggttccatctgtatg 779
```

RESULT 7

```
US-08-637-759-37
; Sequence 37, Application US/08637759
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
```

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: rpms 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: DNA sequence of VGC II from centre to left
; ORGANISM: hand end
; US-08-637-759-37
```

```
Query Match 96.4%; Score 750.6; DB 10; Length 13417;
Best Local Similarity 98.5%; Pred. No. 1.2e-178;
Matches 767; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
Qy 1 atggcacacacaggtaaatgagtggtcttattgcatctggctgtggtctttatttcgaccatta 60
Db 9573 ATGGCACACACAGGTAATGAGTGGCTTATTGCATTGGCTGTGGCTTTATTTCGACCATGTG 9632
Qy 61 agccttttcttacttcccctattaaaaagtgccagtttagggcgcgcttcttcaagt 120
Db 9633 AGCCTTCTCTTATTTACTTCCCTTATTTAAAGTGCAGTTTAAAGGCGGCACCTTTTACGT 9692
Qy 121 aatggcggtctattgctacattccattcccatattaccatattaccacagcagagatt 180
Db 9693 AATGGCGGTCTTATGTCTACTTACCCTTCCGATATTACCAATCATTTACCAGCAGAGATT 9752
Qy 181 atgagtcattatggtaaaagattacacagttggttagggtagtcacgcggagagtgattatt 240
Db 9753 ATGATGCATATTGGTAAAGATTACAGTTGGTTAGGTTAGTACACTGGAGAGGTGATTATT 9812
Qy 241 ggttttcttaattgggtttttgagcggttccctctttggcgcggttgatagtcgggggttt 300
Db 9813 GGTTTTCAATTTGGGTTTGTGCGCGGTTCCCTTTTGGCGCGTTTCATATGATGCGGGGTTT 9872
Qy 301 ctgcttgatactttacgtggcgcgacaatgggtgacgattatcaattctacaatacgaagct 360
Db 9873 CTGCTTGATACTTTTACGTGGCGCGACAAATGGGTAGCATATTCAATTTCAACAAGAGCT 9932
Qy 361 gaaacctcacctttttggcttgccttttcagccagtttttctgtgtgtattttcttataagc 420
Db 9933 GAAACCTCACCTTTTGGCTTGTCTTTACGCCAGTCTTGTGTGTTATTTCCTTTATNAGC 9992
Qy 421 ggcggcgatggagtttatattaaacattctctgtatgagtcataatcaattaccaccagg 480
Db 9993 GCGCGCATGGAGTTTATATAAATCTCTGTATGATCATATCAATATATTTTACCACACGG 10052
Qy 481 cgtactttattattgacggcgcaatttttaaaataatatccagcagagtggaagaagctt 540
Db 10053 CGTACTTTTATTTATTTGACCCAGCAATTTTAAATAATATATCCAGGCAGAGTGGAGAACGCTT 10112
Qy 541 tatcaattatgtctgtctctctctctgcctgcataatataatgtagtgattagccagctcg 600
Db 10113 TATCAATTATGTATCAGCTTCTCTCTCTGCTGCCAATAATATGATGATGATGATGATG 10172
Qy 601 gctttaggtcttttaaacggtcgcgcaacaatgaatgattgtttttctctcaatgcgcg 660
Db 10173 GCTTTAGGTCTTTTAAATCGGTCGGCACAACAATGAATGATGCTGTTTTTCTTCTCAATGCG 10232
Qy 661 ctcaaaagatatattggttctactgacgctcctcgatctcattcccttatgctcttcacac 720
```

Qy	181	atgatgcatttgtaaagattacagtttggttaggttagtcaccgagaggtgattatt	240
Db	9753	ATGATGCATATTGGTAAAGATTACAGTTGGTTAGGTTAGTCACTGGAGAGTGATTATT	9812
Qy	241	ggtttttaattgggttttggcgcgcggttccttlttgggcggttgatatggcgggggttt	300
Db	9813	GGTTTTTCAATTGGGTTTTTGGCGGCGGTTCCCTTTTGGCGCGTTGGATATGCGCGGGTTT	9872
Qy	301	ctgcttgatactttacgtggcgcgacaaagggtacgatatccaattctacaatagaagct	360
Db	9873	CTGCTGTACTTACGTGGCGGCGCAATGGGTACGATATTCCAATTCTACAAATAGAACGT	9932
Qy	361	gaacctcoacttttggctgtcttttcagccagtttttggtgtgtattttctttataagc	420
Db	9933	GAACCTTCACTTTTTGGCTGTCTTTTCAGCCAGTCTCTGTGTGTTATTCTTTATAAGC	9992
Qy	421	ggcgcaatggagtttatataaacattctgtatgagtcataatattttaccaccacggg	480
Db	9993	GGCGGCATGGAGTTTATTTAAACATCTGTATGAGTCATATCAATATTTTACCACCAAGG	1005
Qy	481	cgtactttattttgaccggcaattttttaaataatacccaggcagagtggaagacgctt	540
Db	10053	CGTACTTTATTATTGACCAGCAATTTTTTAAATATATCCAGGCAGAGTGGAGAACGCTT	1011
Qy	541	tatacaattatgtgcagtttctctctctgcataataatgtatggtattagccgactg	600
Db	10113	TATCAATTATGTATAGCTCTCTCTCTCGTCGCAATAATATGTATGTATAGCGATCTG	1017
Qy	601	gccttaggtcttttaactcggtcggcacacaacttgaatgtttttctctcaatcgcg	660
Db	10173	GCCTTAGGCTTTTAAATCGGTGCGGCACAACAATTTGAATGTGTTTTTCTTCTCAATCGCG	1023
Qy	661	ctcaaaagtatatggtgtctactgcgctctgcatactctatcccttatgctcttcaac	720
Db	10233	CTCAAAAGTATATTGTTCTACTGAGC-YCCTGATCTCATTCCTTATGCTTCTCATCAC	1029
Qy	721	tatttggttgaagcgataaaattttatatattctcaaaagactggtttccactgtatg	779
Db	10292	TATTTGGTTGAAGCGATAAAATTTTATATTATCTPAAAGACTGGTTTCATCTGTATG	10350

TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 13417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: DNA sequence of VCC II from centre to left
ORGANISM: hand end
US-08-871-355-37

Query Match 96.4%; Score 750.6; DB 12; Length 13417;
Best Local Similarity 98.5%; Pred. No. 1.2e-178;
Matches 767; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 atggcacacaggaatagtgcttattgcattgctgcttatttcacacacatta 60
|||||
DB 9573 ATGGCACACAGGTAATGAGTGGCTTATTCGATTGGCTGGCTTTTATTCGACCATG 9632
QY 61 agcctttctttattacttccctattataaaagtgagtggttagggcgctctttacgt 120
|||||
DB 9633 AGCCTTCTTTATTACTTCCCTATTAAAGTGGCAGTTTAGGGCGCGCATTTCAGT 9692
QY 121 aatggcgcttattgcaacttaacttccctattcccatattcccaattttaccagagaagatt 180
|||||
DB 9693 AATGGCGTGGCTTATGTCATTTACCTTTCCGATATTTACCAATATTTACCAGCAGAAAGATT 9752
QY 181 atgacatattgtaagattacagtttaggttaggttagtcacgagagagagattatt 240
|||||
DB 9753 ATGATGCATATTCGTAAGATTACAGTTGCTTAGGGTTAGCTACTGGAGAGGTGATTATT 9812
QY 241 ggttttttaattgggttttggcggttcccttttggcggttggatgaggggggttt 300
|||||
DB 9813 GGTTTTTCATTTGGGTTTGTGGCGGGTTCCTTTTGGGCCGTTTGATATGGCGGGGTTT 9872
QY 301 ctgctgacttacttagtgccgcgacaaatgggtacgatatcattcaatataagaagct 360
|||||
DB 9873 CTGCTTGATACTTTACGTGGCGGCACAAATGGGTACGATATTTCAATATTCATAGAAAGCT 9932
QY 361 gaaacctcacttttggctgttttccagccagtttttgggtgttttattttcttataagc 420
|||||
DB 9933 GAAACCTCACTTTTGGCTTGTCTTTTCCAGCCAGTCTTGTGTGTTATTCTTTTATAAGC 9992
QY 421 ggcggcatggagttattataaacattctctgtatgagtcatacaataattaccacacagg 480
|||||
DB 9993 GCGCGCATGGAGTTTATTTAAACATTTCTGTATGAGTCATATCAATATTTACCACACAGG 10052
QY 481 cgtacttattatttaccggcgaatttttataaatacatatccagcagagtgagagacgtt 540
|||||
DB 10053 CGTACTTTTATTATTGACCAAGCAATTTTAAATATATATATCAGCGAGAGTGGAGAACCTT 10112
QY 541 tatcaattatgttcagtttctctcttccctggccataataatgtatggattagccgatctg 600
|||||
DB 10113 TATCAATTATGATACAGTCTCTCTTCCCTGCCATAATATGATGGTATTAGCCGATCTG 10172
QY 601 gctttagggttttataatcgttcggcgaacaaattgaatgtgttttcttctcctaagccg 660
|||||
DB 10173 GCTTTAGGTCTTTTAAATCGTCGGCAACAAATTTGAATGTGTTTCTTCTCAATGCCG 10232
QY 661 ctcaaaagattattggttctactgaagctctctctcattcccttattgctcttctcacc 720
|||||
DB 10233 CTCAAAAGTATATTGGTTCTTACTGAGC-YCCTGATCTCATTCCTCTTATGCTCTTCATCAC 10291
QY 721 tatttggtgaagcgaataattttattattatctaaagagctggtttcccatctgtatg 779
|||||
DB 10292 TATTGTGTTGAACCGGATAAATTTTATATTATCTAAAGACTGTTTCCATCTGTATG 10350

RESULT 10
US-08-637-759-15/c
Sequence 15, Application US/08637759
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: fpm 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 297 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Partial sequence of Salmonella typhimurium
ORGANISM: virulence gene
US-08-637-759-15

Query Match 26.6%; Score 207.4; DB 10; Length 297;
Best Local Similarity 97.2%; Pred. No. 6.8e-42;
Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atggcacacaggaatagtgcttattgcattgctgcttggctttttatttcgaccatta 60
|||||
DB 224 ATGGCACACAGGTAATGAGTGGCTTATTCGATTGGCTGGCTTTTATTCGACCATG 165
QY 61 agcctttctttattacttcccttataaaagtgagtgagtgagggcgctctttacgt 120
|||||
DB 164 AGCCTTCTTTATTACTTCCCTTATTTAAAGTGGCAGTTTAGGGCGGCACCTTTTACGT 105
QY 121 aatggcgcttattgctacttacccttcccatattaccaattaccacagagaagatt 180
|||||
DB 104 AATGGCGTGTATTATGTCTACCTTACCTTTCCGATATTTACCAATCTTTTACCAGCAGAAGATT 45
QY 181 atgacatattgtaagattacagttggttaggtt 217
|||||
DB 44 ATGATGCATATTGGTAAGATTACAGTTGGGTAGGTT 8

RESULT 11
US-08-637-759A-15/c

```

; Sequence 15, Application US/08637759A
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759A
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Partial sequence of Salmonella typhimurium
; ORGANISM: virulence gene
; JS-08-637-759A-15

```

Query Match	26.5%	Score 207.4	DB 10	Length 297
Best Local Similarity	97.2%	Pred. No. 6.8e-42		
Matches 211	Conservative 0	Mismatches 6	Indels 0	Gaps 0
Qy 1	atggcacacaggttaaatgaatggcgttattgcattggctgttgcttttatttcgaccattta 60			
Db 224	ATGGCACACAGGTAATAGTGGCTTATTGCATTGGCTGGCTTTATTTCGACCAATTG 165			
Qy 61	agccttttttattacttccccctattaaaaagtcgcagtttagggcgcgctcttttaacgt 120			
Db 164	AGCCTTTTCTTTATTACTTCCTTATTAAAAGTGGCAGTTTAGGGCGCGCACATTTACGT 105			
Qy 121	aatggcgtgttatgtcaacttaoacttccccatattaccaatcatttaccagcagaagatt 180			
Db 104	AA'TGGCGGTGCTTATGTCACTTACCTTCCGATATTACCAATCATTTTACCAGCAAGATT 45			
Qy 181	atgatgcattatggtaaagattacagtttggttagggt 217			
Db 44	ATGATGCATATGGTAAAGATTACAGTGGGTAGGTT 8			

RESULT 12
US-08-871-355-15/c
; Sequence 15, Application US/08871355
; GENERAL INFORMATION:

```

1  APPLICANT: David William Holden
2  TITLE OF INVENTION: Identification of Genes
3  NUMBER OF SEQUENCES: 38
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: Patrea L. Pabst
6  STREET: 2800 One Atlantic Center
7  STREET: 1201 West Peachtree Street
8  CITY: Atlanta
9  STATE: Georgia
10 COUNTRY: USA
11 ZIP: 30309-3450
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: PatentIn Release #1.0, Version #1.30
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/871,355
20 FILING DATE: 09-JUN-1997
21 CLASSIFICATION: 435
22
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: PCT/GB95/02875
25 FILING DATE: 11-DEC-1995
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Pabst, Patrea L.
28 REGISTRATION NUMBER: 31,284
29 REFERENCE/DOCKET NUMBER: RPNs 101 CON
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (404) 873-8794
32 TELEFAX: (404) 873-8795
33 INFORMATION FOR SEQ ID NO: 15:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 297 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: double
38 TOPOLOGY: linear
39 MOLECULE TYPE: DNA (genomic)
40 HYPOTHETICAL: NO
41 ANTI-SENSE: NO
42 ORIGINAL SOURCE:
43 ORGANISM: Partial sequence of Salmonella typhimurium
44 ORGANISM: virulence gene
45 US-08-871-355-15

```

Query Match	26.6%	Score	207.4	DB	12	Length	297
Best Local Similarity	97.2%	Pred. No.	6.8e-42				
Matches	211	Conservative	0	Mismatches	6	Indels	0
QY	1	atggcaacaacagtgaaatgagtggtattatgcattggtgtggtgtttatttcgaccattta	60				
Db	224	ATGGCAACAACAGTAAATGAGTGGCTTATTCGATTGGCTGGCTTTATTTCGACCAATTG	165				
QY	61	agcctttcttttactctccctattaaaagatggcgagtttaggggcgcgtctttttacgt	120				
Db	164	AGCCTTTCTTTATTACTTCCCTTATTAAGAGTGGCGAGTTTAGGGCGGCACATTTACGT	105				
QY	121	aatggcgtgcttatgtcacttacctttcccatattaccaatcatttaccagcagaagatt	180				
Db	104	AATGGCGTGCATTATGTCACTTACCTTTCCGATATTACCAATCATTTACCAAGCAAGAAT	45				
QY	181	atgatgcataatggtaazagattcacagttggttgggt	217				
Db	44	ATGATGCATATTGGTAAGAAATTACAGTTGGGTAGGTT	8				

RESULT 13
US-09-453-702B-206/c
; Sequence 206, Application US/09453702B
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; ; Valerie
; ; Perna, Nicole T.
;


```

; Plunkett, Guy
; Welch, Rod
;
; TITLE OF INVENTION: Novel Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43360
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-453-702B-206

Query Match 15.1%; Score 118; DB 18; Length 43360;
Best Local Similarity 52.0%; Pred. No. 1.2e-18;
Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 208 tgggttaggttagtcaccgagaggtgattatgggttttttaattgggtttgtgcggcg 267
DB 35984 TTGACAGGAATAGCGCTAAAGAAATAAGTATTGGTTTTTCATTGGTTATCATTTACT 35925

QY 268 gtcccttttggcggtgatagtcgggggtttctgtgatactttacgtggcgcgaca 327
DB 35924 ATTCTTTTGGGCAATAGATCGCGCTGACAGAGATTATTGATACCTTAAGAGGCTCAACA 35865

QY 328 atgggttagcatctcaattctacaatagaagctgaaacctcaactttttggctgtctttc 387
DB 35864 ATATCTTCATTTTAAACCGTCCATAGTATGATGATTCATCTTCTATCACTGGCGTTATCTTG 35805

QY 388 agccagctttgtgtgtattttctttataagcggcgcgatggagtttatataaacatt 447
DB 35804 TACCAATTTATCTCTGTGATCTTTGTTATTTCATGGTGGGATACAAAGCATCTAGATAAG 35745

QY 448 ctgtatagtcatacaatatttaccacagggcggtacittattattttgacggcgcaattt 507
DB 35744 CTATATTTATCTCTACAGATATTACCATTACAAGCCGATATTGCAATTCATCTGCTTTA 35685

QY 508 ttaaaatataccagcagagtggaagacgctttatcaattatgtgtcagtttctctt 567
DB 35684 ATAGATTTTTTGTCTCTAGGGATTCTATTATTAAACTGATGTATCATTTTCAGT 35625

QY 568 cctggcataaatgtatggtatagccgatctggcttttaggtcttttaaatcggtcgcca 627
DB 35624 CCCATGATTATCGGTATATCTTATGCTGATATGGGGTTTGGGTTTCTTAAACAAACAGCA 35565
```

```

QY 628 caacaattgaatgtgtttttctctcaatccgcgtcaaaaagtatattgtgttactgacg 687
DB 35564 CTTCAGCTAAACGTATTACATTATCATCTGCCAGTAAAGATTGTAGCAATCTTTATA 35505

QY 688 ctctgatctcattcccttattgctcttcatt 717
DB 35504 TTGCTATTAGTTATTCATGTTTTTTCCTGAT 35475

RESULT 14
US-09-453-702B-261/C
; Sequence 261, Application US/09453702B
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
;
; TITLE OF INVENTION: Novel Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 261:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 261:
US-09-453-702B-261

Query Match 15.1%; Score 118; DB 18; Length 45325;
Best Local Similarity 52.0%; Pred. No. 1.2e-18;
Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 208 tgggttaggttagtcaccgagaggtgattatgggttttttaattgggtttgtgcggcg 267
DB 36886 TTGACAGGAATAGCGCTAAAGAAATAAGTATTGGTTTTTCATTGGTTATCATTTACT 36827

QY 268 gtcccttttggcggtgatagtcgggggtttctgtgatactttacgtggcgcgaca 327
DB 36826 ATTCTTTTGGGCAATAGATCGCGCTGACAGAGATTATTGATACCTTAAGAGGCTCAACA 36767

QY 328 atgggttagcatctcaattctacaatagaagctgaaacctcaactttttggctgtctttc 387
DB 36766 ATATCTTCATTTTAAACCGTCCATAGTATGATGATTCATCTTCTATCACTGGCGTTATCTTG 36707
```

QY 388 agccaggttttgggtgttatttcttattataagcgcgccgaggtttattataaacatt 447
 Db 36706 TACCAATTTATCTGTCATCTTTGTTTATTCATGTCGGATACAAAGCATCCTAGATAAG 36647
 QY 448 ctgtatgagtcataatattaccaccagggcgtaacttattattttgaccggaattt 507
 Db 36646 CTATATTTATCCTACGAGATATTACCAATTACAAGCCGATATTCGATTCAAATCGTGTTA 36587
 QY 508 ttaaaatataatccaggcgagagtggaacgcttttataatattatgtcagtttctctt 567
 Db 36586 ATAGATTTTTGTTTCTCTATGGGATTCATTTATTAACATGATGTTATCATTTTCAGTT 36527
 QY 568 cctgccataatattatgttattagccgagtcgtttagctctttttaaatacggcgca 627
 Db 36526 CCCATGATTTATCGTATATCTTATGATATGGGTTTGGGTTCTTAAACAAACAGCA 36467
 QY 628 caacaattgaatgtttttcttctcaatccgcgtcaaaagtattgttctactgacg 687
 Db 36466 CCTCAGCTAAACGTTTACATTCATCTGCCAGTAAAAAGTTTGTAGCAATCTTTATA 36407
 QY 688 ctctgatctcattcccttattgctcttcat 717
 Db 36406 TTGCTATTAGTTTATTCATGTTTTTCTTGAT 36377

RESULT 15

US-09-409-800B-1/c
 ; Sequence 1, Application US/09409800B
 ; GENERAL INFORMATION:
 ; APPLICANT: Blattner, Frederick R.
 ; APPLICANT: Burland, Valerie
 ; APPLICANT: Rose, Debra J.
 ; APPLICANT: Mayhew, George F.
 ; APPLICANT: Perna, Nicole
 ; APPLICANT: Perry, Robert D.
 ; APPLICANT: Straley, Susan C.
 ; APPLICANT: Fetherston, Jacqueline D.
 ; APPLICANT: Lindler, Luther E.
 ; APPLICANT: Plano, Gregory V.
 ; TITLE OF INVENTION: Plasmid DNA From Yersinia Pestis
 ; FILE REFERENCE: 960296.95939
 ; CURRENT APPLICATION NUMBER: US/09/409,800B
 ; CURRENT FILING DATE: 1999-09-30
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 70559
 ; TYPE: DNA
 ; ORGANISM: Yersinia pestis
 US-09-409-800B-1

Query Match 9.8% Score 76; DB 18; Length 70559;
 Best Local Similarity 48.2%; Pred. No. 5.6e-08;
 Matches 245; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
 QY 193 ggttaagattacagttggttaggttagtcaccggagaggtgattattggttttttaatt 252
 Db 27477 GTTGATGCGTTTACGTGTCCTTATCGGCAAGAGATCATCTGGGGTTATTGATT 27418
 QY 253 ggttttggcggttcccttttggccgttgatagcggggtttctgcttgatact 312
 Db 27417 GGGTTTGGCCACCATCTCTTCTGGGCTTAGAATCCGAGGATTTATTGTAGATAAC 27358
 QY 313 ttagtggtcgcaatgggttagatattcaattctcaatagaagctgaacctcatt 372
 Db 27357 CAAAGAGTCCGGATGGGATCTCTACTCAATCTCGACTGTAGTCAAACTAGTCCG 27298
 QY 373 ttgggttgttttcaagccagttttgtgtttatttttttataagcgcgcatggag 432
 Db 27297 ACTGGTTTACTTTTGACTCAACGTTAATNACATTTTCTTCAGCGGAGGGGCTTCCTC 27238
 QY 433 ttatatataaacattctgtatgagtcataatattaccaccagggggtactttatta 492

Db 27237 TCTTTTGGCTTTCAGCCCTCTTTTTCACAGCTATGTAAATTTGGCCGGTGGCCAGCTTTTTCCT 27178
 QY 493 tttagaccggcaatttttataatataatccaggcgagagtggaagagcttttatacaattatgt 552
 Db 27177 GCAGTTAGTGAACAGTGGGTTTCATTTCTTTTATAACCAATTCAGCCAGATATTAATC 27118
 QY 553 gtcagtttct---ctcttccctgccataatattatgtatgtattagcagctggtgcttaggt 609
 Db 27117 GCCGCTGTATTGGGCTGCTCTCTTTTACTCATCGCTATGTTTTTAGCTGAATTTGGACTTCA 27058
 QY 610 cttttaaatacggtcggcgcaacaatttgaatgtgttttttcttctcaatggcgctcaaaagt 669
 Db 27057 CTCATCAGTCGTTTTCGCGCTTCTTTAAACGCTTTTGTGCTGCTATGCCGATAAAAAAGC 26998
 QY 670 atattggttctactgacgcctctgactct 697
 Db 26997 GCGATAGCAAGCCCTGTTGTTGGTTATCT 26970

Search completed: August 2, 2002, 03:50:14
 Job time: 7579 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sv model

Run on: August 2, 2002, 02:16:20 ; Search time 4280.15 Seconds
(without alignments)
395.143 Million cell updates/sec

Title: US-09-809-524B-1
Perfect score: 779
Sequence: 1 atggcacacaggttaataga.....gactgggttcactctgtatg 779

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1359166 seqs, 1085540162 residues

Total number of hits satisfying chosen parameters: 2718332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	118	15.1	43360	US-10-114-170-206	Sequence 206, App
C 2	118	15.1	45325	US-10-114-170-261	Sequence 261, App
C 3	53.6	6.9	16950	US-10-114-170-166	Sequence 166, App
C 4	37.2	4.8	14147	US-10-172-086-51	Sequence 51, Appl
C 5	36.6	4.7	283	US-09-539-800C-12433	Sequence 12433, A
C 6	36.6	4.7	301	US-09-539-800C-14817	Sequence 14817, A
C 7	36.6	4.7	22713	US-10-105-299-12732	Sequence 12732, A
C 8	36	4.6	11445	US-06-391-781-423	Sequence 423, App
C 9	35.2	4.5	621	US-10-027-632-313321	Sequence 313321, A
C 10	35.2	4.5	621	US-10-027-632-313322	Sequence 313322, A
C 11	35.2	4.5	2460	US-10-104-047-770	Sequence 770, App
C 12	35	4.5	536	US-09-721-544-1708	Sequence 1708, App
C 13	34.8	4.5	168575	US-10-178-194-1	Sequence 1, Appl
C 14	34.4	4.4	509	US-10-027-632-87445	Sequence 87445, A
C 15	34.2	4.4	611	US-10-027-632-179972	Sequence 179972, A
C 16	34.2	4.4	647	US-10-027-632-187203	Sequence 187203, A
C 17	34.2	4.4	2484	US-09-629-469A-17966	Sequence 17966, A
C 18	34.2	4.4	2713	US-09-629-469A-16837	Sequence 16837, A
C 19	34	4.4	520	US-09-918-995-19426	Sequence 19426, A
C 20	34	4.4	98686	US-10-189-267-20	Sequence 20, Appl
C 21	33.8	4.3	465	US-09-918-995-10260	Sequence 10260, A
C 22	33.8	4.3	1295	US-10-144-132-33	Sequence 33, Appl
C 23	33.8	4.3	113585	US-10-188-470-12	Sequence 12, Appl
C 24	33.6	4.3	311	US-09-991-936-1266	Sequence 1266, App
C 25	33.6	4.3	585	US-10-027-632-193565	Sequence 193565, A

RESULT 1
US-10-114-170-206/c
; Sequence 206, Application US/10114170
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Nicole T.
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: Novel Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43360
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-10-114-170-206

Sequence 235108,
Sequence 235109,
Sequence 235110,
Sequence 235111,
Sequence 235112,
Sequence 248133,
Sequence 248133,
Sequence 224256,
Sequence 18755, A
Sequence 18756, A
Sequence 2656, Ap
Sequence 12126, A
Sequence 12126, A
Sequence 223574,
Sequence 7220, Ap
Sequence 34696, A
Sequence 34697, A
Sequence 145891,
Sequence 145891,
Sequence 27598, A
Sequence 163, App
Sequence 163, App
Sequence 129, App

ALIGNMENTS

```
Query Match      15.1%; Score 118; DB 6; Length 43360;
Best Local Similarity 52.0%; Pred. No. 1.8e-19;
Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 208 tggtaggttagtcacccgagaggtgatttgggttttttaattgggttttgcggcg 267
Db 35984 TTGACAGGAATAGCGCTAAAGAAATAGTATTGGTTTTTTCATTTGGTTATCATTTACT 35925

QY 268 gtcccttttgggcccgttgatagcggggtttcttgcgttatacttttacggcgagaca 327
Db 35924 ATTCCTTTTGGGCAATAGATGGCGCTGGACAGATTATTGATACCTCTAAGAGGCTCAACA 35865

QY 328 atgggtagcagatattcaattacagaagctgaacccctcaactttttggcttgccttttc 387
Db 35864 ATATCTTCAATTTTAAACCGCTCCATAAGTGATTTCATCTTCTATCATCTGCGTTATCTTG 35805

QY 388 agccagttttgtgttattttctttataaagcggcggaatggagtttatataaacatt 447
Db 35804 TACCAATTTATCTGTGATCTTTGTATTTCATGTGGGATACAAAGCATCTCTAGATAAG 35745

QY 448 ctgtatgagtcataatcaattattaccaccagggcgctactttattttacccggcaattt 507
Db 35744 CTATATTTCCTACGAGATATTACCATTTACAAGCCGATATTGCAATTCATCGTGTCTTA 35685

QY 508 ttaaaatataccagcagagtgaggaacgctttatcaattatgtcagtttctctctt 567
Db 35684 ATAGATTTTGTGTTTCTCTATGGGATTCATTTATTAACCTGATGTATTATTTTCAGTT 35625

QY 568 cctgccataatgtatgtattagcagatcgtgcttttaggtttttaaatacggcgaca 627
Db 35624 CCCATGATTATCGGTATATCTTATGTATGGGTTTGGTTTCTTAACAAACAGCA 35565

QY 628 caacaattgaatgttttcttctcaatgcgcgtcaaaagctatattggttctactgacg 687
Db 35564 CCTCAGCTAAACGATTCACATTATCAGTCGCCAGTAAAGTTTGATAGCAATCTTTATA 35505

QY 688 ctccgtatcctatcccttatgctcttcat 717
Db 35504 TTGCTATTAGTTATTTCATGTTTTTCTCTGAT 35475

RESULT 2
US-10-114-170-261/G
; Sequence 261, Application US/10114170
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: Novel Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinkney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
```

```
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 261:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 261:
US-10-114-170-261
```

```
Query Match      15.1%; Score 118; DB 6; Length 45325;
Best Local Similarity 52.0%; Pred. No. 1.9e-19;
Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 208 tggtaggttagtcacccgagaggtgatttgggttttttaattgggttttgcggcg 267
Db 36886 TTGACAGGAATAGCGCTAAAGAAATAGTATTGGTTTTTTCATTTGGTTATCATTTACT 36827

QY 268 gtcccttttgggcccgttgatagcggggtttcttgcgttatacttttacggcgagaca 327
Db 36826 ATTCCTTTTGGGCAATAGATGGCGCTGGACAGATTATTGATACCTCTAAGAGGCTCAACA 36767

QY 328 atgggtagcagatattcaattattaccaccagggcgctactttttggcttgccttttc 387
Db 36766 ATATCTTCAATTTTAAACCGCTCCATAAGTGATTTCATCTTCTATCAGTCGGCTTAATCTTG 36707

QY 388 agccagttttgtgttattttctttataaagcggcggaatggagtttatataaacatt 447
Db 36706 TACCAATTTATCTGTGATCTTTGTATTTCATGTGGGATTCATTTATTAACCTGATGTATTATTTTCAGTT 36527

QY 448 ctgtatgagtcataatcaattattaccaccagggcgctactttatttttacggcgcaattt 507
Db 36646 CTATATTTCCTACGAGATATTACCATTTACAAGCCGATATTGGCATTCATTCGTCGCTTTA 36587

QY 508 ttaaaatataccagcagagtgaggaacgctttatcaattatgtcagtttctctctt 567
Db 36586 ATAGATTTTGTGTTTCTCTATGGGATTCATTTATTAACCTGATGTATTATTTTCAGTT 36527

QY 568 cctgccataatgtatgtattagcagatcgtgcttttaggtttttaaatacggcgaca 627
Db 36526 CCCATGATTATCGGTATATCTTATGTATGGGTTTGGGTTTCTTAACAAACAGCA 36467

QY 628 caacaattgaatgttttcttctcaatgcgcgtcaaaagctatattggttctactgacg 687
Db 36466 CCTCAGCTAAACGATTCACATTATCAGTCGCCAGTAAAGTTTGATAGCAATCTTTATA 36407

QY 688 ctccgtatcctatcccttatgctcttcat 717
Db 36406 TTGCTATTAGTTATTTCATGTTTTTCTCTGAT 36377
```

```
RESULT 3
US-10-114-170-166/G
; Sequence 166, Application US/10114170
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: Novel Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Quarles & Brady
STREET: 1 South Plinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA: US/10/114,170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:
LENGTH: 16950
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 166:
US-10-114-170-166

Query Match 6.9%; Score 53.6; DB 6; Length 16950;
Best Local Similarity 49.6%; Pred. No. 0.0015;
Matches 137; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
QY 216 gtagtcacggagaggtgattatggttttttaattggttttgcggcggttccctt 275
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5518 GTTAGCTGCAAGAGAAATTATCATAGGCGTGTCATTGCTGCTTATATAGCATCTCCATT 5459
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 276 ttgggcccgttgatggcggtgtctctgctgatactttacgtggcgcgacaatgggtac 335
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5458 TTGGATTTTCCTTGCTATAGTAGTACCTTATCGATAACCAACGCGGTGCAACGTTAAGTAG 5399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 336 gatattcaatttcacaatagaagctgaacacctcacttttgggttgccttttcagccagtt 395
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5398 TACTTTGGATCCGCTACAGGTGCGATACCTCGGAATTAGCAAGATTATTAAACCTTTT 5339
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 396 ttigtgttatatttttttaagcgcgcatggagtttatataaacattctgtatga 455
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5338 CTCTGACGCGGTTTATCTTACTACGCTGGATTGAACCTTTATTCTGGAAACTCTTTAGCA 5279
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 456 gtcatatcaatatttaccacacggcgctactttatt 491
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5278 AAGTTATAATTTATGGCTTCAGGAAGTTTAAATTT 5243
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 4
US-10-172-086-51
; Sequence 51, Application US/10172086
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; OF PROSTATE TUMORS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13

; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 51
; LENGTH: 14147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-172-086-51
Query Match 4.8%; Score 37.2; DB 7; Length 14147;
Best Local Similarity 46.6%; Pred. No. 18; Mismatches 173; Indels 1; Gaps 1;
Matches 152; Conservative 0;
QY 173 agaagattatgatcatatttgtaaagattacagatttggttggttagtgtagtcacggagag 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2290 ataaaatgtggagatatttggtatataatggtttatatttttttttttttttttttt 2349
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 233 tgattattggttttttaattgggttttgcggcggttcccttttggcggttgatgg 292
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2350 atattttatttattggttttttttttaaaatagttatttagtttttttttttttt 2409
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 293 cgggggttctgctgatactttacgtggcgcgacaatgggacgatactca-attctaca 351
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2410 ttgattttttatgatagataataaaaatttttttttttttttttttttttttttttt 2469
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 352 atagaagctgaacacctcacttttgggttgccttttcagccagtttttgggttttttttc 411
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2470 tgatttaatt 2599
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 412 tttataagcgcgcatggagtt 471
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2530 ttttaataatttgatt 2589
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 472 ccaccaggcgctactttatttga 497
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2590 gtggaaagggttt 2615
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 5
US-09-539-800C-12433/c
; Sequence 12433, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999

```

; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 12433
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01079302
US-09-539-800C-12433

```

[illegible]

```

RESULT 6
US/95-539-800C-14817/c
: Sequence 14817, Application US/09539800C
: GENERAL INFORMATION:
: APPLICANT: Sellhauer, Jeffrey J.
: APPLICANT: Deleceane, Angelo M.
: APPLICANT: Stuart, Susan L.
: APPLICANT: Stuve, Laura G.
: APPLICANT: Mullahy, Sara J.
: TITLE OF INVENTION: POLYNUCLEOTIDES OF C
: FILE REFERENCE: PD-1023 CIP
: CURRENT APPLICATION NUMBER: US/09/539,800
: CURRENT FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: 08/521,383
: PRIOR FILING DATE: August 16, 1995
: PRIOR APPLICATION NUMBER: 08/271,217
: PRIOR FILING DATE: June 27, 1994
: PRIOR APPLICATION NUMBER: 08/334,881
: PRIOR FILING DATE: November 4, 1994
: PRIOR APPLICATION NUMBER: 08/943,978
: PRIOR FILING DATE: October 3, 1997
: PRIOR APPLICATION NUMBER: 60/028,732
: PRIOR FILING DATE: October 4, 1996
: PRIOR APPLICATION NUMBER: 08/943,979
: PRIOR FILING DATE: October 4, 1997
: PRIOR APPLICATION NUMBER: 60/027,782
: PRIOR FILING DATE: October 4, 1996
: PRIOR APPLICATION NUMBER: 08/993,774
: PRIOR FILING DATE: December 18, 1997
: PRIOR APPLICATION NUMBER: 60/034,975
: PRIOR FILING DATE: December 20, 1996
: PRIOR APPLICATION NUMBER: 09/250,003
: PRIOR FILING DATE: February 10, 1999
: PRIOR APPLICATION NUMBER: 60/074,364
: PRIOR FILING DATE: February 12, 1998
: PRIOR APPLICATION NUMBER: 09/452,747
: PRIOR FILING DATE: December 1, 1999

```

```

; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 14817
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01140786
US-09-539-800C-14817

```

	Query Match	4.7%	Score 36.6;	DB 5;	Length 301;
	Best Local Similarity	50.3%;	Pred. No. 12;		
	Matches 90;	Conservative 0;	Mismatches 89;	Indels 0;	Gaps 0;
QY	236	ttatgggtttttaaattgggttttggcggggttcccttttggggccgttgatggcgg	295		
Db	277	TTCTTTCTTCTTATTAAAGGCTTGTGTGCTGGGTTCCTAGTTGCAACTTAGATAGAAAA	218		
QY	296	ggttctgcttgatacatttcacgtggcgcgacaaagggtacgatatccaattctcaaatag	355		
Db	217	GATTCTTGTGAGAGCTTTTCTTCTATCCAAATTAGCTTCAGTTTCCATTTTCACATCAT	158		
QY	356	aagctgaagacctcacttttggcttgcttttccagccagtttttggtggtatttcttt	414		
Db	157	TACCACTAGTGTATCTGTGAGAGTTATGTCTCTGTGTCCTTTTCTCTACTACTTTT	99		

```

RESULT      7
US-10-105-299-12732/c
; Sequence 12732, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12732
; LENGTH: 22713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-12732

```

Query Match	4.7%	Score 36.6	DB 7	Length 22713
Best Local Similarity	55.0%	Pred. No. 28		
Matches 72	Conservative 0	Mismatches 59	Indels 0	Gaps 0
625	gcacacaaattgaatggtgtttttcttcctcaatgccgcctcaaaagtatatatggtctctactg	684		
6261	GCTCAATAAATACATATTATTATTTCTCATCACATCTCTCACTATATTTTCTTTTACTT	6202		
685	acgctcctgatactcatccctctatgctcttcacacatttggttgaaagcgataaatatt	744		
6201	CTCATCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	6142		
745	tatatattctatct	755		
6141	TTTTTTTTTTCT	6131		

RESULT 8
US-60-391-781-423/c
; Sequence 423, Application US/60391781
; GENERAL INFORMATION:
; APPLICANT: Laurie, Cathy C.
; TITLE OF INVENTION: Nucleic Acid Mol

; TITLE OF INVENTION: Nucleic Acid Molecules Associated With Oil in Plants

```

; FILE REFERENCE: 38-77(52900)B
; CURRENT APPLICATION NUMBER: US/60/391,781
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/365,301
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 3034
; SEQ ID NO 423
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: zea mays
US-60-391-781-423

Query Match          4.5%; Score 36; DB 8; Length 1145;
Best Local Similarity 55.6%; Pred. No. 22;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 626 cacacaaatgaagtgttttcttcataagccgctcaaaagtagtatgttgtactga 685
    ||||| || | ||||| || | ||||| || | ||||| || | ||||| || | |||||
DB 873 CACAATTTTACACTGTGCTCCGTCACATACATATGCCAGCATATATTGAAGCTCATGC 814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 686 cgcctcgtatcatctccattgatctcttcattcacatttggtggaaagcgataatttt 745
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 813 AGCTCCTATTCTCAAAACCAGTTTTCAGCGCCCAAATGTTTGCTTGAAGAATAAAGTTC 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 746 atat 749
    |||
DB 753 AGAT 750

RESULT 9
US-10-027-632-313321
; Sequence 313321, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313322
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(621)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-313322

Query Match          4.5%; Score 35.2; DB 7; Length 621;
Best Local Similarity 48.1%; Pred. No. 31;
Matches 100; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 443 acattctgtatgagtcacatcaatatattaccacaggcgcttattattgtgccggc 502
DB 234 agatttttaaatgctctttcttcattagtaattattagtaacattatttagtatactgat 293
QY 503 aatttttaaataatatccaggcagtgaggagaagcgtttacaattatgtgtcagtttct 562
DB 294 aaataataataataatagtgatgactgtatctatctattattaggtttttttctataaatt 353
QY 563 ctcttctgcctaataatagtgatgattagcgcgactctggcttttaggtctttttaaatcgat 622
DB 354 atattattaatttataaagtttaattccccccaattctgtctattcttctattctttttct 413
QY 623 cggcacacaacaaatgaatgtgtttttctt 650
DB 414 caatctcctgattgcatttattttt 441

US-10-027-632-313322
; Sequence 313322, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313322
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(621)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-313322

Query Match          4.5%; Score 35.2; DB 7; Length 621;
Best Local Similarity 48.1%; Pred. No. 31;
Matches 100; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 443 acattctgtatgagtcacatcaatatattaccacaggcgcttattattgtgccggc 502
DB 234 agatttttaaatgctctttcttcattagtaattattagtaacattatttagtatactgat 293
QY 503 aatttttaaataatatccaggcagtgaggagaagcgtttacaattatgtgtcagtttct 562
DB 294 aaataataataataatagtgatgactgtatctatctattattaggtttttttctataaatt 353
QY 563 ctcttctgcctaataatagtgatgattagcgcgactctggcttttaggtctttttaaatcgat 622
DB 354 atattattaatttataaagtttaattccccccaattctgtctattcttctattctttttct 413
QY 623 cggcacacaacaaatgaatgtgtttttctt 650
DB 414 caatctcctgattgcatttattttt 441

```

```
RESULT 11
US-10-104-047-770
; Sequence 770, Application US/10104047
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770
; LENGTH: 2460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-770

Query Match          4.5%; Score 35.2; DB 7; Length 2460;
Best Local Similarity 46.4%; Pred. No. 40;
Matches 115; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 233 tgattattggttttttaattggttttgcggcggttccttttggccgttgatatgg 292
      ||||| ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 tgaagccaaattctcaatttttttttgatgttgttcttttcaggagtgagaaatct 2151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 293 cgggggtttctgtgatactttacgtgcgcgacaatgggtacgatatattcaattctacaa 352
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2152 atgattggtgtgttcatttccagctctctctgtagcattctcgttgaattcttaac 2211
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 353 tagaagctgaacctcaatttttggctgttgcagcagtttttgcgttattttct 412
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2212 tagagatgtacattttctaaacctgttggttatccatgtctgtctgtacatatga 2271
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 413 ttataagcggcgcatgtagtttatattataaacattctgtatgagtcattcaattttac 472
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2272 ccatgtttcttgacatggaaataccagtcagtggttcctatgtattgtctatatattgcaa 2331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 473 caccagg 480
      |||||
Db 2332 accaagag 2339
      |||||

RESULT 12
US-09-721-544-1708
; Sequence 1708, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jones, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Trach, Joe
; APPLICANT: Tran, Lien
```

```
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fei
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721,544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1708
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(536)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-544-1708

Query Match          4.5%; Score 35; DB 5; Length 536;
Best Local Similarity 54.2%; Pred. No. 33;
Matches 71; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 371 tttttggttcttttcagccagtttttgggtgtgttttttttataagcggcgcatgg 430
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 tctctggtagcttttaagtcgaagtttttgaagtgttttaagtttgaagaataa 446
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 431 agtttatataaacattctgtatgctatcatcatatattaccacaggcggtactttat 490
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 447 actcattgacaaacacttagaataatccaaagtaatagtatttttggcaccagtttt 506
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 491 tatttgaccgg 501
      |||||
Db 507 gtttgtatgg 517
      |||||

RESULT 13
US-10-178-194-1/c
; Sequence 1, Application US/10178194
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/10/178,194
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/426,290
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21181)..(21403)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95252)..(95430)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101753)..(101996)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (110324)..(110439)
; FEATURE:
```



```
; NAME/KEY: CDS
; LOCATION: (124058)....(124278)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127009)....(127130)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128910)....(129139)
US-10-178-194-1

Query Match      4.4%; Score 34.8; DB 7; Length 168575;
Best Local Similarity 52.8%; Pred. No. 11e+02;
Matches 75; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 427 atggagtttattaaacattctgtatgagtcataatcaccagcaggtggagaacgtttataca 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18358 ATTCAAGTTATAAATACGAGAAATGAAATATTTTCAAGATTTTATCCTGATACCTGCT 18399

Qy 487 ttattattgacggcaatttttataatataatccagcagagtgaggagaacgtttataca 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18298 TTATTTATGACAAGCTGATTTATCTTGATCTTACAAATATTCCTTCTCTATCTA 18239

Qy 547 ttatgtgcagttctctcttc 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18238 TGATTTCTCTGGCTTCTTC 18217

RESULT 14
US-10-027-632-87445
; Sequence 87445, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 87445
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-87445

Query Match      4.4%; Score 34.4; DB 7; Length 509;
Best Local Similarity 46.2%; Pred. No. 46;
Matches 110; Conservative 1; Mismatches 127; Indels 0; Gaps 0;

Qy 336 gatattcaattctacaatagagctgaaacacctcttttggctgttttcagccagtt 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 51 gattttaagtgaaatgagtatacaaaaacacattttctcacaagcattataatgaaat 110

Qy 396 ttgtgtgtattttcttataaaggcgaggtttatataaaccattctgtatga 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 gttgaaygaacatttttttgaggactgtgtatatagtttttgcttaaaagctgtgtccc 170
```

```
Qy 456 gtcatacaataattaccaccagggcgtaactttattatttaccggcaatttttaaaata 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 aagaacttattgacattgaagactgtctatctgtcccttacttccatatcttc 230

Qy 516 tatccaggcagagtggaagacgtttatcaattatgtgtcagtttctcttctccgccc 573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 tctgtaagtgaactctcacaaactatttctccagtcctcaaatatctgtcgaatatccgcc 288

RESULT 15
US-10-027-632-179972
; Sequence 179972, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 179972
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-179972

Query Match      4.4%; Score 34.2; DB 7; Length 611;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 111 tctttacgttaatggcgcttatgtcaccttacccttcccatattaccacattaccaca 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 tatatcccatatttagccttaatttttccaattttctcttattatgtttccattttatca 264

Qy 171 gcagaagatttatgatgcataattggtaagattaca 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 gaataagattattcttttatgtaggcaaaattttcca 299

Search completed: August 2, 2002, 05:02:21
Job time: 9961 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2002, 00:37:05 ; Search time 1811.32 Seconds

(without alignments)

8999.940 Million cell updates/sec

Title: US-09-809-524B-1

Perfect score: 779

Sequence: 1 atgcacacaggttaaatga.....gactggttccatctgtatg 779

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sv.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sta.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

SUMMARIES

Result No.	Score	Match	Length	ID	Query	Description
------------	-------	-------	--------	----	-------	-------------

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

1	779	100.0	779	6	AX254677	AX254677 Sequence
2	763	97.9	779	6	AX254678	AX254678 Sequence
3	763	97.9	5484	1	STYSCORS	X99944 S.typhimuri
4	763	97.9	27863	1	AE008761	AE008761 Salmonella
5	750.6	96.4	13417	6	A51693	A51693 Sequence 37
6	750.6	96.4	13417	6	AX001627	AX001627 Sequence
7	207.4	26.6	297	6	A51671	A51671 Sequence 15
8	207.4	26.6	297	6	AX001605	AX001605 Sequence
9	180.8	23.2	288050	1	AJ414141	AJ414141 Yersinia
10	119.6	15.4	35624	1	AF022236	AF022236 Escherich
11	118	15.1	11760	1	AE005597	AE005597 Escherich
12	118	15.1	45325	1	AF071034	AF071034 Escherich
13	118	15.1	280800	1	AP002566	AP002566 Escherich
14	113.8	14.6	37889	1	AF200363	AF200363 Escherich
15	113.8	14.6	60073	1	ECO277443	AJ277443 Escherich
16	113.8	14.6	60354	1	AF453441	AF453441 Escherich
17	99.2	12.7	42001	1	AF311901	AF311901 Citrobact
18	76	9.8	5752	1	YEPYSC	L25667 Yersinia ps
19	76	9.8	70305	1	YPCD1	AL117189 Yersinia
20	76	9.8	70504	1	AF053946	AF053946 Yersinia
21	76	9.8	70559	1	AF074612	AF074612 Yersinia
22	73.8	9.5	67720	1	AF336309	AF336309 Yersinia
23	70.6	9.1	110000	2	LMFLCHR32_02	Continuation (3 of
24	67.8	8.7	7874	1	SHRPMYSH	D13663 Shigella fl
25	67.8	8.7	20323	1	SHEORF	D50801 Shigella so
26	67.8	8.7	213494	1	SFPMR100	AL391753 Shigella
27	67.8	8.7	221618	1	AF386526	AF386526 Shigella
28	67.8	8.7	221851	1	AF348706	AF348706 Shigella
29	66.6	8.5	303249	1	AP001515	AP001515 Bacillus
30	66.4	8.5	69673	1	AF102990	AF102990 Yersinia
31	60.2	7.7	801	6	AX028438	AX028438 Sequence
32	59.2	7.6	11958	1	AE004596	AE004596 Pseudomon
33	53.8	6.9	13008	1	AE007715	AE007715 Clostridi
34	53.6	6.9	10875	1	AE005515	AE005515 Escherich
35	52.4	6.7	3350	1	AF172245	AF172245 Bordetell
36	52.2	6.7	184	6	A51672	A51672 Sequence 16
37	52.2	6.7	184	6	AX001606	AX001606 Sequence
38	50	6.4	1641	1	AF306650	AF306650 Sodalis g
39	49.6	6.4	9613	1	AB052736	AB052736 Escherich
40	47.2	6.1	7218	6	I66494	I66494 Sequence 14
41	44.6	5.7	1673	8	AF284928	AF284928 Carex fra
42	44	5.6	124508	30	AL139017	AL139017 Human DNA
43	44	5.6	124975	2	AC104455	AC104455 Homo sapl
44	44	5.6	197605	2	AC026905	AC026905 Homo sapl
45	43.8	5.6	117847	2	AC095778	AC095778 Rattus no

ALIGNMENTS

RESULT	1	AX254677	779 bp	DNA	linear	PAT 10-OCT-2001
LOCUS	AX254677	Sequence 1 from Patent WO0170247.				
DEFINITION	AX254677					
ACCESSION	AX254677.1	GI:16074344				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						

Salmonella enterica subsp. enterica serovar Dublin.

Salmonella enterica subsp. enterica serovar Dublin

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

Salmonella

1 (bases 1 to 779)

Lowery, D.E. and Kennedy, M.J.

Salmonella vaccine materials and methods

Patent: WO 0170247-A 1 27-SEP-2001;

Pharmacia & Upjohn (US)

Location/Qualifiers

source

i. 779

/organism="Salmonella enterica subsp. enterica serovar

Dublin"

/db_xref="taxon:98360"

BASE COUNT 185 a 143 c 161 g 290 t

ORIGIN

KEYWORDS

pathogenicity island; secretion system apparatus; ssaQ gene; ssaR gene; ssaS gene; ssaT gene; ssaU gene.

SOURCE

Salmonella typhimurium.

ORGANISM

Salmonella typhimurium
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.

REFERENCE

1 (bases 1 to 5484)

AUTHORS

Hensel,M.

TITLE

Direct Submission
Submitted (09-AUG-1996) M. Hensel, Max von Pettenkofer-Institut, Lehrstuhl fuer Bakteriologie, Pettenkofer Strasse 9a, 80336 Munich, FRG

JOURNAL

79175536 Location/Qualifiers

FEATURES

1..3484

REFERENCE

2 (bases 1 to 5484)

AUTHORS

Hensel,M., Shea,J.E., Baumber,A.J., Gleeson,C., Blattner,P. and Holden,D.W.

TITLE

Analysis of the boundaries of Salmonella pathogenicity island 2 and

JOURNAL

the corresponding chromosomal region of Escherichia coli K-12

MEDLINE

J. Bacteriol. 179 (4), 1105-1111 (1997)

FEATURES

79175536 Location/Qualifiers

source

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

FEATURES

1..3484

Db 2232 GCTTTAGGCTTTTAAATCGGTCGGCACAAATTAAGATGTGTTTTCTCTCAATGCCG 2291

QY 661 ctcaaaagtatatgttctactgagctcctgagctcctgagctcctccttatgctcttcatac 720
 |||||

Db 2292 CTTCAAAAGTATATGTTCTACTGAGCCCTCTGATCTCATTCCTTATGCTCTTCATCAC 2351
 |||||

QY 721 tatttggtgaagcagataaaatttatatttatctctaaagactggtttccatctgtatg 779
 |||||

Db 2352 TATTGCTGAAAGCGATAAATTTATATATTATCTCTAAAGACTGGTTCCATCTGTATG 2410
 |||||

RESULT 4

AE008761

LOCUS

DEFINITION

Salmonella typhimurium LT2, section 69 of 224 of the complete genome.

ACCESSION

AE008761 AE006468

VERSION

AE008761.1 GI:16419908

KEYWORDS

SOURCE

ORGANISM

Salmonella typhimurium LT2.

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.

REFERENCE

AUTHORS

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grew, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R. and Wilson, R.K.

TITLE

Complete genome sequence of Salmonella enterica serovar Typhimurium LT2

JOURNAL

Nature 413 (6858), 852-856 (2001)

PUBMED

11677509

REFERENCE

2 (bases 1 to 27863)

AUTHORS

The Salmonella typhimurium Genome Sequencing project.

TITLE

Direct Submission

JOURNAL

Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA

COMMENT

Supported by NIH grant 5U 01 AI43283

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and GeneMark; <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc; <http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and ReguonDB; http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

Location/Qualifiers

1..27863

/organism="Salmonella typhimurium LT2"

/strain="LT2; SGSC 1432; ATCC 700720"

/db_xref="ATCC:700720"

/db_xref="taxon:99287"

70..75

/gene="orf70"

/note="Putative RBS for orf70; RegulonDB:STMSIH001626"

70..301

/gene="orf70"

FEATURES

source

1..301

/note="S. typhimurium hypothetical 7.9 Kda protein in ssrB-ttrR intergenic region (ORF70). (SW:YDHz_SALTY)"

/codon_start=1

/transl_table=11

/product="putative cytoplasmic protein"

/protein_id="AAL20312.1"

/db_xref="GI:16419909"

/translation="MTKPTKDELYREMCRVGVGVVLEMDLGOEPKYIIVIAGVLRTA LANQRIQSALEKQAMETVINALARS"

complement(397..1367)

/gene="orf319"

/note="STM1389"

complement(397..1356)

/gene="orf319"

/note="S. typhimurium ORF 319 (gi|4456867)"

/codon_start=1

/transl_table=11

/product="putative inner membrane protein"

/protein_id="AAL20313.1"

/db_xref="GI:16419910"

/translation="MNNKPVVGISGCTGATVDFDGHKRMDFVMSLAPWVAYKPTC PEVAIGLPVPALRLIQTTCGDIRMYSHAPDDVTERMAADRELTIGELAGFI VCAKSPCGMERVRLYDEKNGRKGAGTGLFTAAAMDKYPLWFIEDGRLHDPVLREN FIARIPALHELNALRAQGLSRHSLAFHSRYKIQLLAHQAGYREIGPPFVARLHWD DLDFFVREKLMATLRHSPASRNKNTNLMHIOGYFPRALNSQRAELREVILGYRAG RLDPILAPLTLKHLYLAEPDDPYLRSONYFEPYPTDLGLRLAIT"

complement(1362..1367)

/gene="orf319"

/note="Putative RBS for orf319; RegulonDB:STMSIH001627"

complement(1529..2257)

/gene="orf242"

/note="STM1390"

complement(1529..2257)

/gene="orf242"

/note="S. typhimurium ORF 242 (gi|4456866)"

/codon_start=1

/transl_table=11

/product="putative regulatory proteins, merr family"

/protein_id="AAL20314.1"

/db_xref="GI:16419911"

/translation="MSYSGEFARLGCINAATLRAWORRYGLLKPORTDGGHLYSDD IYDCGREYPAELVTHLRPLRSKVSAPLPAVMTLREILDGIIATFTSCLEGGRRAP GNAFTSGNLSNDHCEIWELETITRTIGQELRLNLVSPVPVLAPELFAQRKWFVTTGK LTAGQKKLAQRNVAASLEVITL"

complement(2436..3079)

/gene="ssrB"

/note="STM1391"

complement(2436..3074)

/gene="ssrB"

/note="S. typhimurium SsrB (gi|2765824)"

/codon_start=1

/transl_table=11

/product="Secretion system regulator: transcriptional activator, homologous with degu/uvrY/bvgA"

/protein_id="AAL20315.1"

/db_xref="GI:16419912"

/translation="MKEYKILLVDDHEIINGINNALLPWPHFKEIVHKNGLEYVNA CCAYEPDILDLSPFGINGLDIIPOLHQRPANMLIVYTAQQEYMTIKTLAAGANG VYKSSQOVLIALAALQAVANKRYIDPTLNREAILAELNADTTHNOLLTLRROVLKL IDEGYTNHIGISEKLHISIKTVETHRMMNRKLQVHKVTELLNCARRMLLEY"

complement(3074..3079)

/gene="ssrB"

/note="Putative RBS for ssrB; RegulonDB:STMSIH001628"

complement(3105..5874)

/gene="ssrA"

/note="STM1392"

complement(3105..5867)

/gene="ssrA"

/EC_number="2.7.3.-"

```

/EC_number="2.7.1.40"
/note="S. typhimurium SpIR (gi|1498305)"
/codon_start=1
/transl_table=11
/product="Secretion system regulator:Sensor component"
/protein_id="AAL20316.1"
/db_xref="GI:16419913"
/translation="MNLLNKNTLQTSIVIRLTFLELLTIIWLLSVLTAAVYSWQ
RKHIEDLSVSEMNIVLSNQRSEAEERDAKNIMYCISLATEIHNDIPPEVSRHLS
VGPNCNTLNGEKHRTFLQSSDIENSFGRSPFLNKKNEISLLSDSPSDSYLQSP
LTRKSFYTPHAGYWESEYINGKQHASVAVAQQQGVFEVTVKLPDLITKSHLP
DDSSIRVMDQNNHLFPFSYIPQIRKTQLENVTLHDGQQIIPGFILRTLTHGQWSL
LTVYVGNHRILKILQIIPFTLTLVMTSAFCWLLHRSKLARWVVDVINKTA
TAPLSRLPDAKRLDILOSQIAGAFNOLDLTQOVNDLENKVAERTQALNEAKKRAERA
NKKRSIHLTVISHELRTPMNCVCAIELQTLPLNBOGLADTARNCTLSLAIINN
LLDPSRIESGHTLMEETALLPILDQAMQTIQGPASCKKLSLRTFVGQHVPLFHTF
SIRQLVNLGNHNAFTGGRKLTVRKHEQLIFPVSDSGRGIEIQOQSQITAF
LADPNSGTGIGITIASSPHQQMGNLTKSPVGTCSVLPLQEQEYPOPIKGT
LSAPFCRLHQACIGIRGEPPHQOANLNAELLYFGSKLVLQAOELTCTPNPPIKT
LLPFWQIQLLVLDHNPNDIRDTIGKMLVSLAQHVTIAASSNEALT SQOQRFDLVLVDI
RMPETIDGECVRLVWHDEPNNDIPDKMVFALSASVATEDIHRCKNGIHHVITKPTVLA
TLARYISAAEQVQLLNLELQEQDPSRCSALLATDDMVINSKIFQSILDLALDIENAV
SAGEKIDQLIHTLKGCLQIQGTVELVCYVDIENRVKMKIIALEELTDLRQKIRMF
KNRIT"
complement(5869..5874)
/gene="ssrA"
/note="Putative RBS for ssrA; RegulonDB:STMS1H001629"
/size=6669
/gene="ssab"
/note="STM1393"
/size=6669
/gene="ssac"
/note="S. typhimurium SpIC (gi|1498306)"
/codon_start=1
/transl_table=11
/product="Secretion system apparatus"
/protein_id="AAL20317.1"
/db_xref="GI:16419914"
/translation="MSEEGFMVLAVKGIPLQDTRAGNSRSWMTIDGHPARGEIPS
EAFSISFLNDESLEPRCLAYVPLLLAAHPDHDVYAIQLTADGGWLNYYTTSSSSE
LIAIEIKHLALTCLIKNVIRNHHKLSGGV"
/size=8164
/gene="ssac"
/note="STM1394"
/size=6670
/gene="ssac"
/note="Putative RBS for ssac; RegulonDB:STMS1H001630"
/size=8164
/gene="ssac"
/note="S. typhimurium SpIA (gi|1498307)"
/codon_start=1
/transl_table=11
/product="Secretion system apparatus"
/protein_id="AAL20318.1"
/db_xref="GI:16419915"
/translation="MVNKKRILILLFILNTAKSDELSWRGNDFTLYARQMPLEAVLH
LSSEYDITAITSPITAFHLSQKIPPGPPDILNNLAAQYDLITWFGSMPLVYPASL
LKQVITNTILSTGFIHYLRSONILSPGCEVKEITGTRAVEVSGVPSCLTRISOLA
SVLDNALIKRDSASVSISYTLKATAMDQYQVRDOSVVPVGVVLRMSKTSVPT
STNGNSPATQALPMFADPRONAVIPDYAANWAGYRKILITELDQOQHIETSVKII
DVNAGDINQLGDWGTAVSLGKKIAFNGLNDGGAGSGFSTVSDTNFVVRNLNALEK
SSQATVLSQVSVPTLNNIQAVLDNNITVYTKLOGEVAKLESITGSLLRVTRPLKLD
NOTKIMLNLTNDQOQSDETQSDPPEVONSIAOSQSLLAGQSLLAGGFGQCKQOI
HSQNKIPLLGDIPIVVGHLFRNDTTQVHVSIFRLFIKASVYVNNNGISHG"
/size=9356
/gene="ssad"
/note="STM1395"

```

Query Match 97.9%; Score 763; DB 1; Length 27863;
Best Local Similarity 98.7%; Pred. No. 6.4e-174;
Matches 769; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db	25924	ATGGCACAACAGGTAAATGAGTGGCTTATTGCATTTGGCTGGGCTTTTATTTCGACCAATTG	25983
Qy	61	agccttctcttactctccctctataaaagtcgagctttaggggcccgcctctttacgt	120
Db	25984	AGCCTTCCTTTATTACTTCCTCTTATTAAGAGTGGCAGTTTATAGGGCGCGACTTTTACGT	26043
Qy	121	aatggcgctgtattgtcacttcccttcccatattaccaatcattaccagcagaagaatt	180
Db	26044	AATGGCGTGCCTTATGTCTACATTACCTTTCCGATATTACCAATCATTTACCAGCACAAGATT	26103
Qy	181	atgatgcatttgttaaaagattacagtttggttaagggttagtcaocggagagaagtgattatt	240
Db	26104	ATGATGCATATTGGTAAGATTTACAGTTGGTTAGGTTAGTCTACTGGAGAGGTGATTATT	26163
Qy	241	ggtttttcaattgggttttctgcgcgcggttccctcttctggcgcttgatcgagcggggttt	300
Db	26164	GGTTTTTCAATTGGGTTTTGTGCGCGGTTCCCTTTTGGGCGGTTGATATGGCGGGGTTT	26223
Qy	301	ctgtctgatactttacgtggcgcgacaatgggttcagcatattcaattctacaatagaagct	360
Db	26224	CTGCTTGATACTTTACGTGGCGGACAAATGGTACGATATTTCAATTCTACAATAAGAGCT	26283
Qy	361	gaacctcaacttttggcttgccttcttcagccaggttttggtagtgattttctttataaagc	420
Db	26284	GAACCTCTACATTTTGGCTTCCTTTTCAGCCAGCTTCTTGTGTGTTATTTCTTTATAGC	26343
Qy	421	ggcgctgaggtatttatataaacattctgtatgagtcataatcttaccaccaggg	480
Db	26344	GGCGCATGGAGTTTATATTAAACATTCGTATGAGTCATATCAATATTTACCACGAGG	26403
Qy	481	cgactctattatttgaccgggcaattttttaaataatccagggcagagtgagagaacgctt	540
Db	26404	CGTACTTTTATTATTGACCAGCAATTTTAAATATATCCAGGCAGAGTGGAGAACGCTT	26463
Qy	541	tatcaattatgtctgaatttctctctctccatcaatatgtatggttatagccgagctg	600
Db	26464	TATCAATTATGTATCAGCTTCTCTCTCTGCCATAATATGTATGGTATTAGCCGATCTG	26523
Qy	601	gctttagctcttttaaatcggctggcgacaacaaattgaatggtttttctctcaatgccc	660
Db	26524	GCTTTAGCTCTTTTAAATCGGTGCGCACAATTTGAATGTGTTTTCTCTCAATGCGG	26583
Qy	661	ctcaaaagtatatggctctactgaagcctctgactctcattcccttatgtctcttcacac	720
Db	26584	CTCAAAAGTATATTGGTTCTACTGAGCTCCTGATCTCATTTCCCTTATGCTCTTCATCAC	26643
Qy	721	tatttgggtgaaagcgataaaatttttatatttatctcaaaagactggtttcccatcgtatg	779
Db	26644	TATTTGGTTGAAGCGGATAAATTTTATATTATCTAAAAAGACTGTTTCCCATCTGTATG	26702
RESULT	5		
LOCUS	A51693	Sequence 37 from Patent WO9617951.	linear PAT 10-MAR-1997
DEFINITION	A51693		
ACCESSION	A51693		
VERSION	A51693.1	GI:2304497	
KEYWORDS		unidentified.	
SOURCE		unidentified.	
ORGANISM		unclassified.	
REFERENCE		1 (bases 1 to 13417)	
AUTHORS		Holden,D.W.	
TITLE		IDENTIFICATION OF GENES	
JOURNAL		Patent: WO 9617951-A 37 13-JUN-1996;	
COMMENT		RPMS TECHNOLOGY LTD (GB)	
FEATURES		Other publication AU 4121996 960626.	
source		Location/Qualifiers	
		1. .13417	
		/organism="unidentified"	
		/db_xref="taxon:32644"	
BASE COUNT	3580 a	2816 c	3301 q 3718 t 2 others

Qy 1 atggcacacaggtaaatgagtggcttatcgcatctggctgtggcttttatctgaccatta 60

ORIGIN

[illegible]

JOURNAL	environment
Patient: EP 0889120-A 37 07-JAN-1999;	
IMP COLLEGE INNOVATIONS LTD (GB)	
FEATURES	Location/Qualifiers
source	1. .J3417
	/organism="unidentified"
	/db_xref="taxon:32644"
BASE COUNT	3580 a 2816 c 3301 g 3718 t
ORIGIN	2 others
Query Match	96.4%; Score 750.6; DB 6; Length 13417;
Best Local Similarity	98.5%; Pred. No. 6.6e-171;
Matches 767; Conservative	1; Mismatches 10; Indels 1; Gaps
QY 1	atggcacaacagtgaaatgagtggtctattgcattggctgtgtgctctttatgttcgaccattta 60
Db	9573 ATGGCAACACAGTAATGAGTGGCTTATGCAATGGCTGTGGCTTTTATGTACCACTTG 9632
QY 61	acccctttctttattactccctattaaaagtggcagtttaggggcgcgtctctttacgt 120
Db	9633 AGCCCTTCTTTTATTACTTCCCTTATTTAAAAGTGGCAGTTTAGGGCCGCACTTTTACGT 9692
QY 121	aatggcgctgtctatgtcacttacccttcccatattaccaatcattaccagcagaagatt 180
Db	9693 AATGGCGTGTATTATGTCACTTACCTTTCGGATATTACCAANTCATTTACCAGCAGAAGATT 9752
QY 181	atgatgcatacttggttaagattcacagttgggttagtgtagtcacccggagaggtgtattt 240
Db	9753 ATGATGTCATATTGGTAAAGATTACAGTTGGTTAGGTAGTACCTGGAGAGGTGATTATT 9812
QY 241	ggtttttcaattgggttttggcgccggttcccttttggccgttgatattggcgggggttt 300
Db	9813 GGTTTTTCAAATTGGGTTTGTGGCGCGGTTCCTTTTGGCGGTTTGATATGGCGGGGTTT 9872
QY 301	ctgcttgatactttacgttggcgcacaatgggtacgatattcaattcacatagaagct 360
Db	9873 CTGCTTGATACTTTACGTGGCGGCAATGGGTACGATATTCAATTCACAAAGAGCT 9932
QY 361	gaaacctcaccttttggctgtctttcagccagtttttgtgtgtattttctttataaagc 420
Db	9933 GAAACCTCACCTTTTGGCTTGCTTTTCAGCCAGTTCCTGTGTGTATTCTTTATAAGC 9992
QY 421	ggcgcatgaggtttatataaacattctgtatgagtcataatcaatattaccaccagg 480
Db	9993 GCGGCATGGAGTTATATTAAACATTCTGTATGAGTCATATCAATATTACCACAGG 10052
QY 481	cgtactttattattgaccgcgaatttttaaaatatatccaggcagagtgagaaacgctt 540
Db	10053 CGTACTTTATTATTGACCAGCAATTTTAAATATATCCAGGCAGAGTGAGAAACGCTT 10112
QY 541	tatcaattatgttgcagtttctctctctgccataatgtatggtattagccgactgtg 600
Db	10113 TATCAANTATGTATCAGCTTCTCTCTCCGCCATATATATGTATGGTATTAGCCGATCTG 10172
QY 601	gctttaagttcttttaaaatcggctggcacaacaaattgaattgttttctctctcaatgc 660
Db	10173 GCTTTAGGTCTTTTAAATCGGTTCGGTCGGCAACAATTTGAATGTGTTTCTTCTCAATGCC 10232
QY 661	ctcaaaagtattggttctactagcctcgtatcattcccttatgctctctcatcac 720
Db	10233 CTCAAAAGATATATGGTCTTACTGACG - YCCTGATCTCATTCCTCTATGCTCTCTCATCAC 10291
QY 721	tatttggtgaaagcgataaattttatatttatctaaagactgggttccatctctgtatg 779
Db	10292 TATTTGTTGAAAGCGATAAAATTTTATATTATCTAAAGACTGGTTTCCATCTGTATG 10350
RESULT 7	
AS1671/c	
LOCUS	A51671
DEFINITION	Sequence 15 from Patent WO9617951.
ACCESSION	A51671
	linear PAT 10-MAR-19


```
VERSION A51671.1 GI:2304475
SOURCE .
ORGANISM Salmonella typhimurium.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
1 (bases 1 to 297)
AUTHORS Holden,D.W.
TITLE IDENTIFICATION OF GENES
JOURNAL Patent: WO 9617951-A 15 JUN-1996;
COMMENT RIMS TECHNOLOGY LTD (GB)
FEATURES Other publication AU 4121996 960626.
source Location/Qualifiers
1..297
/organism="Salmonella typhimurium"
/db_xref="taxon:602"
BASE COUNT 99 a 67 c 55 g 76 t
ORIGIN
Query Match 26.6%; Score 207.4; DB 6; Length 297;
Best Local Similarity 97.2%; Pred. No. 8.2e-40;
Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 atggcacaacaggtaaatgagtgcttattgcattgctgctgcttattgacaccatta 60
DB 224 ATGGCACAACAGGTAAGTGGCTTATTGCAATGGCTGCTGCTTATTTCGACCAATTG 165
QY 61 agcctttcttattactccctattataaagtgccagtgtagggcgctcttttaccgt 120
DB 164 AGCCTTCTTATTACTCTCCCTTATTAAAGTGGCAGTTAGGGCGGCACCTTTTACGT 105
QY 121 aatggcgctgattgctacattcccttcccatattaccaatcattaccagcagaagatt 180
DB 104 AATGGCGTCTTATGTCACCTACCTTTCCGATATTACCAATCATTTACCAGCAGAAGATT 45
QY 181 atgatgcattgtaagattacagttggttaggt 217
DB 44 ATGATGCATATTGTTAAGATTACAGTTGGGTAGGTT 8
RESULT 9
AJ414141 288050 bp DNA linear BCT 04-OCT-2001
LOCUS Yersinia pestis strain C092 complete genome; segment 1/20.
DEFINITION AJ414141 AL590842
ACCESSION AJ414141.1 GI:15978115
VERSION
KEYWORDS
SOURCE
ORGANISM Yersinia pestis.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.
1 (bases 1 to 288050)
REFERENCE Parkhill,J., Wren,B.W., Thomson,N.R., Titball,R.W., Holden,M.T.G.,
AUTHORS Prentice,M.B., Sebaihia,M., James,K.D., Churcher,C., Mungall,K.L.,
Baker,S., Basham,D., Bentley,S.D., Brooks,K., Cerdeno-Tarraga,A.M.,
Chillingworth,T., Cronin,A., Davies,R.M., Davis,P., Dougan,G.,
Fellayshew,A.V., Moule,S., Oyston,P.C.F., Quail,M., Rutherford,K.,
Simmonds,M., Skelton,J., Stevens,K., Whithead,S. and Barrall,B.G.
TITLE Genome sequence of Yersinia pestis, the causative agent of plague
JOURNAL Nature 413 (6855), 523-527 (2001)
MEDLINE 21470413
REFERENCE 2 (bases 1 to 288050)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) Submitted on behalf of the Yersinia
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT Details of Y. pestis sequencing at the Sanger Centre are available
on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/Y_pestis/).
FEATURES
source Location/Qualifiers
1..288050
/organism="Yersinia pestis"
/strain="C092"
/db_xref="taxon:632"
/notes="biovar: Orientalis"
complement(271..711)
/gene="YPO0001"
complement(271..711)
/gene="YPO0001"
/notes="Similar to the FMN-binding domain of many
flavoproteins e.g. Salmonella typhimurium sulfit
reductase [NADPH] flavoprotein alpha-component CysJ
SW:CYSL_SALTY (P38039) (598 aa) fasta scores: E():
3.8e-06, 32.0% id in 128 aa. Also highly similar over
entire range to Escherichia coli MIOC protein, the function
of which in a chromosomal context, is unknown
SW:MIOC_ECOLI (P03817) (146 aa) fasta scores: E(): 0,
66.9% id in 145 aa"
/codon_start=1
/transl_table=11
/product="putative flavoprotein"
/protein_id="CAC8868.1"
/db_xref="GI:15978116"
/translation="MADITLISGSLTGSAEYVAFHADKLERAGFSTEILHCGPELDEL
TLNGLWLIIVTSHGAGDLPDNLQPLLEQKPDLSQVRFPAVGLGSSEYDFPCGAI
IKLDDQLIAQAQRLEILSDIVQHEIPEDPAEIVWKDWINLL"
QY 1 atggcacaacaggtaaatgagtgcttattgcattgctgctgcttattgacaccatta 60
DB 224 ATGGCACAACAGGTAAGTGGCTTATTGCAATGGCTGCTGCTTATTTCGACCAATTG 165
QY 61 agcctttcttattactccctattataaagtgccagtgtagggcgctcttttaccgt 120
DB 164 AGCCTTCTTATTACTCTCCCTTATTAAAGTGGCAGTTAGGGCGGCACCTTTTACGT 105
QY 121 aatggcgctgattgctacattcccttcccatattaccaatcattaccagcagaagatt 180
DB 104 AATGGCGTCTTATGTCACCTACCTTTCCGATATTACCAATCATTTACCAGCAGAAGATT 45
QY 181 atgatgcattgtaagattacagttggttaggt 217
DB 44 ATGATGCATATTGTTAAGATTACAGTTGGGTAGGTT 8
RESULT 8
AX001605/c
LOCUS AX001605 297 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 15 from Patent EP0889120.
ACCESSION AX001605
VERSION AX001605.1 GI:7241734
KEYWORDS
SOURCE Salmonella typhimurium.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
1 (bases 1 to 297)
AUTHORS Holden,D.W.
TITLE A micro-organism having reduced adaption to a particular
environment
JOURNAL Patent: EP 0889120-A 15 07-JAN-1999;
IMP COLLEGE INNOVATIONS LTD (GB)
FEATURES Location/Qualifiers
1..297
/organism="Salmonella typhimurium"
/db_xref="taxon:602"
BASE COUNT 99 a 67 c 55 g 76 t
ORIGIN
Query Match 26.6%; Score 207.4; DB 6; Length 297;
Best Local Similarity 97.2%; Pred. No. 8.2e-40;
Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 atggcacaacaggtaaatgagtgcttattgcattgctgctgcttattgacaccatta 60
DB 224 ATGGCACAACAGGTAAGTGGCTTATTGCAATGGCTGCTGCTTATTTCGACCAATTG 165
```

```

gene
complement(804. .1265)
/ gene="asnC"
/ note="YPO0002"
CDS
complement(804. .1265)
/ gene="asnC"
/ note="Similar to the Escherichia coli regulator of Ansa:
      AnC SW:ASNC_ECOLI (P03809) (152 aa) fasta scores: E(): 0,
      84.8% id in 151 aa"
/ codon_start=1
/ transl_table=11
/ product="regulatory protein"
/ protein_id="CAC88869.1"
/ db_xref="GI:15978117"
/ translation="MSEYQIDNLSRSILKALMENARTPYAELAKNLAVSPGTIHVRV
      EMKRGQITACVHNPKQLGDYCCFGITLKSAKDYPSALKKLESLEEVAYIT
      TGHYIFIKVMCKSDALQOVLINKIQITDEIQSTETLISLQNPIMRTIVP"
misc_feature
complement(870. .1175)
/ gene="asnC"
/ note="Pfam match to entry PF01037 ASNC_trans_reg, AnC
      family, score 63.50, E-value 4.4e-15"
misc_feature
complement(1113. .1193)
/ gene="asnC"
/ note="PS00519 Bacterial regulatory proteins, asnC family
      signature."
misc_feature
complement(1131. .1196)
/ gene="asnC"
/ note="Predicted helix-turn-helix motif with score 1363
      (+3.83 SD) at aa 6-27, sequence TPYAELAKNLAVSPGTIHVRV"
gene
1435. .2427
/ gene="asnA"
/ note="YPO0003"
CDS
1435. .2427
/ gene="asnA"
/ EC_number="3.3.1.1"
/ note="Similar to Escherichia coli aspartate-ammonia
      ligase AsnA SW:ASNA_ECOLI (P00963) (330 aa) fasta scores:
      E(): 0, 78.5% id in 330 aa"
/ codon_start=1
/ transl_table=11
/ product="aspartate-ammonia ligase"
/ protein_id="CAC88870.1"
/ db_xref="GI:15978118"
/ translation="MKKQFTQKQQQISFVKSFFSRQLEQOGLLIEVQAPILSRVDCGT
      QDNLSGSEKAVQVVKVLPDSTFVHSLAKWKRTKGRFDFGADQGVYTHKALRPD
      EDRLSAITHSVVDQWDERVYMGDGERNLAYLKVSTVKIYAAIKETEAAISRGVRPF
      LPDHQFTHSESLRFPDLDAKGERAIKELGAVELIGIGKLDQGDHVRAPYD
      DMTSPSAGFSGINGDIIVNPILEDAFELSSMGITRVDALKRQALQDGEDRLLEL
      EWHQSLRGENPQTIGGGIGQSRLVMLLLQKHIGQVQCWVGWGEISEKVDGLL"
gene
complement(2526. .3992)
/ gene="YPO0004"
CDS
complement(2526. .3992)
/ gene="YPO0004"
/ note="Similar to Vibrio cholerae hypothetical protein
      VCA0762 TR:Q8KLI2 (EMBL:AE004404) (481 aa) fasta scores:
      E(): 0, 34.1% id in 496 aa"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="CAC88871.1"
/ db_xref="GI:15978119"
/ translation="MISLATLMLLSISEGELIEEMVVGLLAAPQAIFFKFPRIKR
      ALMKDIPGMQKQKQRIEASVPPGLANESLYQSLLEDSPQFYAHLPDIVAQLQDL
      HSPFATQAKTLVQADLAKNPPGSDLSQTLFLQWRVSLIQITITHHLLQEERQL
      LAELQRLSAGALEITLTNDNAAGRLMDMSQHLQGDYQLLOYGDFLQOQPELI
      RLAEQGRSRAQAPADAPYEPYTVMVROPDSVPEEYSGIHQSDNILLRLPTELVM
      LGMSELEFEYRLLERLLTYRLOQDNQKQTPQSPVSKQNDQPRGPFIVCVDTIS
      GSGMGNEQACAFCLALLRIALADNRCTIMLFATEIHYELSDNGTEQALRFNLQ
      HFGGTDLAACLANKMDEWDADAVIISDFIAQRLPELVKRIKIQQAQAHQR
      FHAVMSYAGKGMIRFDHWRFDTSLSRLIRWRK"
gene
complement(3996. .5549)
/ gene="YPO0005"
CDS
complement(3996. .5549)
/ gene="YPO0005"

```

Query Match 23.2%; Score 180.8; DB 1; Length 288050;
 Best Local Similarity 54.5%; Pred. No. 1.3e-33;

```

/ note="Similar to Escherichia coli hypothetical protein
      YIEN SW:YIEN_ECOLI (P31473) (498 aa) fasta scores: E(): 0,
      67.8% id in 497 aa"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="CAC88872.1"
/ db_xref="GI:15978120"
/ translation="MAQSQAERISRLSHAESGLYERQEAIRLCILLAALSGESVFL
      LGPGIAKSLIARLKFAFRAFAEYILTRFSTPEEVFGPLSQALKKEGRYQRMGTG
      GYLPEAEIWFDEIWKAGPAIINTLLTAINNERFRNGDRSDISPMRLLVATSNELPDA
      DSSLEALYDEMLRIWLDRVOEKFRSLILSRONENHPVAENLSITDRFQHWQPL
      IDKITDPHCFELIFOLRORLSALEHAPYVSDRWKALKALLOASAFSGRDEITPID
      LLLKDCILWHLNLSFKLLQOQLEQLLEEQGQQOSSLMLKLODINSKWLQHQEAOQSDHQ
      ALTVYKSGMFRGAQAALPNLTDSTLTLLQPLNLHLQVNLQVHLQVHQAQAQMLN
      AGGLARAKLNGVGAQSIDAIEDDOLHIITLLDVRQPSSTLSLPGATTSTVPELLAL
      TKLESTLAEQRRLEFSHQPCLFPTSSWLAKTEASLLQVQLOQFOQIQFOQIQQRK
      FQOQKHSCH"
gene
5823. .7691
/ gene="kup"
/ note="YPO0006; trkD"
5823. .7691
/ gene="trkD"
/ note="Similar to Escherichia coli low affinity potassium
      uptake system protein Kup SW:KUP_ECOLI (P30016) (622 aa)
      fasta scores: E(): 0, 84.2% id in 622 aa"
/ codon_start=1
/ transl_table=11
/ product="potassium transport protein"
/ protein_id="CAC88873.1"
/ db_xref="GI:15978121"
/ translation="MSTEHKQYLSAVTLAAIGVVYGDIGTSPLYTLRECFSGHYGFDV
      RPDVFGFLSLIFWMLILVYVKLYVVRADNAGEGGILITLSLAGRNTSSRATSIIL
      VVLGLGGSFGEYVITPAISVMSAIEGIEAIPALDPYIPVPCSTAVLTLLFVIOKH
      GTSGVKLFAFVMDVWFLTLALGLRSITIANPEVLAALNPKAISFVEYKVSFFAL
      GAVLALDVEALYADMGHCFKPIRLAWFTVVLPSLVNLVYFGQALLKNPDAIKNP
      FFLADPMALIPILLATVATIASQAVISGVFSLRQAVRLGYLPMRITHTSEMES
      GOIYIPVNTVLYLVVIIGFSSNLAAYGIATVGTWITSILFCTVAKNWHW
      NRELVLMLVLLIIDIPMSANVLESGWPLSLGLVFWFIITWTKSRFSLLR
      MHEHNSLEAMIASLEKSPVPGTAVMSRAMVIPPALLHNKHLKHVHERVLL
      TMETDDVPYVHNVERTIEQLSTFWRVARYGWRETNPVAFIFRCGLEGSCQKME
      TSFMSHESILITKRPWHLFLRGKLFIALSRNALRAPDQFEIPPNRVIEIGTQVEI"
misc_feature
/ gene="kup"
/ note="one of 12 probable transmembrane helices predicted
      for YPO0006 by TMHMM2.0"
5952. .6020
/ gene="kup"
/ note="one of 12 probable transmembrane helices predicted
      for YPO0006 by TMHMM2.0"
6126. .6194
/ gene="kup"
/ note="one of 12 probable transmembrane helices predicted
      for YPO0006 by TMHMM2.0"
6324. .6392
/ gene="kup"
/ note="one of 12 probable transmembrane helices predicted
      for YPO0006 by TMHMM2.0"
6450. .6518
/ gene="kup"
/ note="one of 12 probable transmembrane helices predicted
      for YPO0006 by TMHMM2.0"
6555. .6623
/ gene="kup"
/ note="one of 12 probable transmembrane helices predicted
      for YPO0006 by TMHMM2.0"

```

[illegible]

```

/codon_start=1
/transl_table=11
/product="Orf1"
/protein_id="AAC38364.1"
/db_xref="GI:2865271"
/translation="MRRLLFMNMTNSHTSPYIQLIEQIAVLQOEAKRLREQEVQSV
EALLGMKEEPLVDTEETEEIPLANI"
4474..4692
/codon_start=1
/transl_table=11
/product="Orf2"
/protein_id="AAC38365.1"
/db_xref="GI:2865272"
/translation="MITTELEDETIKNKEAANVFIEKINDKKNIEHEKMKHPLDKVT
INEAKELLIACDAALRTIEIMRINNK"
4696..5019
/codon_start=1
/transl_table=11
/product="Orf3"
/protein_id="AAC38366.1"
/db_xref="GI:2865273"
/translation="MSIVSQTRNKEELDKKIRSEIEAKKIIAEFDVVKESVNELSEK
AKTPQAAEKLNLKIEGTYGEERKLYDSALSKIEKLETLSPARSKSQSTMNORRN
NRKIV"
5016..5615
/codon_start=1
/transl_table=11
/product="Orf4"
/protein_id="AAC38367.1"
/db_xref="GI:2865274"
/translation="MTIFNKIDYNYKLYNPIMWVDEWGLDTGVNQVNSFRRLRET
SSTRNQLKILROFIODKISGVELSDINKEGFLYOSIGKIRLLALSSALFDIOCPDI
FSRLYRETLREIGYQNVKLSFYWGQCKPEYGEERFCAELIKYGAGNLEWLFADN
PLWTVIYLLPKSGEIKPHTHINDLFLNRLNKLILPYETL"
5560..6255
/codon_start=1
/transl_table=11
/product="Orf5"
/protein_id="AAC38368.1"
/db_xref="GI:2865275"
/translation="MIYFLDLKFKYCLMKLYEFKNIEIDLVLTEIDIEIPDKLQEIIO
SDDIILARKTTEHLLRARRKKSEKELIESRKKIARKMIAMRERIRKNNKIKLDKEVN
QSIKWVDIOAIELVMDIMNKHVLSLTNALSHLDTSSRNWDDLNEVVRLETLSNN
NIVGAKITKNPDILKDPGEANNIQLINDANTPHNKIIIEVEYIRITLDPLEQISILL
NSFKDNLVLSIIQE"
6260..6913
/gene="escr"
6260..6913
/gene="escr"
/codon_start=1
/transl_table=11
/product="Escr"
/protein_id="AAC38369.1"
/db_xref="GI:2865276"
/translation="MSQLMTIGSOPFLIIVFLLSLLPIFVVGTSFLKISIVLGIL
KNALGIQVPPNNALTSVSLITFMTSPITLIQNDNISQEPINYNWDSDFQKVDKI
LSPYRGLEKTEKDNVERFERAAOKKLGNETILKKDSLFIILLPATFTMGQLEAFKIG
FLYLPFIADLLIISNLLALGNMYSPTISIPFKILLIFILVGGWQKLFEFLVVN"
6913..7182
/gene="escs"
6913..7182
/gene="escs"
/codon_start=1
/transl_table=11
/product="Escs"
/protein_id="AAC38370.1"
/db_xref="GI:2865277"
/translation="NDTGYFVQLCVQTFWIIIFILSPTVIAASVIGIISLVQAITQL
QDQLPFLKIIIAVFATLALTYHMGTTIINFSSIIIFMPKVG"
7182..7958
/gene="escr"
7182..7958

```

```

/gene="escr"
/codon_start=1
/transl_table=11
/product="Escr"
/protein_id="AAC38371.1"
/db_xref="GI:2865278"
/translation="MNEIMTVIVSSFYCILLRPLGMFILLIFSTGVLLSNFIRNSIMI
AFTPIIVENYTESEKLPISGIFQLTGVLAEISGFIIGLSFTLLFWAIDAGQIIDT
LRGSTISSINPISDSSSTIGVLYQFISVIFVHGGIOIKLOKLYSLYEILPLOAD
IAFNRLIDFLFLSWDSFIKLMISFVPMIIGFLCDMGEGFLNKTAPQLNVTLSLP
VKSIIAIFILLLVIVHVPDFITANIHSIDIIRSLPSMINE"
7951..8988
/gene="escu"
7951..8988
/gene="escu"
/codon_start=1
/transl_table=11
/product="Escu"
/protein_id="AAC38372.1"
/db_xref="GI:2865279"
/translation="MSEKTEKTPPKLRDLKKKGDVTKSEEVMAAVOSLILFFSFFSLY
CMSFFVDVGLVNTTIDSLARPELYAIREILGAVLNILFLYILPISLIVGVGTTGV
SQGFIPAEKIKPSAOKISVKNMLKNIFSKEIFELLKSVFKLVIILVLFYFMGHSY
ANEFANFTGLNAYQALVYVAFVFLWGLVFLYLLFSVDFEFKHEGLKKWMSKD
EVKREAKDTGNPEIKGERRLHSEIQSGSLANNIKKSTVIVKNPHTHIAICLYKLGE
TLPVLVETGCKDAKALQIKLAELYDIPVIEDIPLARSLYKKNHKQYITDEFPEVA
OLIRIADLDY"
complement(8985..9443)
/codon_start=1
/transl_table=11
/product="rorf3"
/protein_id="AAC38373.1"
/db_xref="GI:2865280"
/translation="MKIILSILISNYCNASDCFITGKAYNIDPLILKATAWNSK

Query Match      15.4%; Score 119.6; DB 1; Length 35624;
Best Local Similarity 52.2%; Pred. No. 8.4e-19;
Matches 266; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

QY 208 tggtaggggttagtcaccggaggggattattggttttttaattgggtttgtggegog 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7383 TTGACAGGATAGTGGCTAAAGAAANTANGATTGGTTTTTTCATTGGGTATTACTTACT 7442
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 gttcccttttgggcgttgataggggggtttctgctgatacttacttgcgcgcgaca 327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7443 ATTCTTTTGGGCAATAGATCGGCTGGCAGATTATTGATCTACTCTAAGAGGTTCAACA 7502
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 328 atgggtacgatattcaattctacaatagaagctgaacacctcacttttggctgtcttc 387
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7503 ATATCTTCAATTTTAAACCGCTCCATAAGTGATTTCATCTTCTATCACTGGCGTTATTGTG 7562
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 agccagttttgtgtattttcttataagcgcgcgatggagttattataaacatt 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7563 TACCAATTATCTCTGTGATCTTGTATTATCATGGTGGGATACAAAGCATCTGGATAAG 7622
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 ctgtatgagtcataatcaattttaccaccaggcgctactttatttattgaccggcaatt 507
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7623 CTATATTTATCCPACGAGATATTACCAATTACAAGCCGATATTGCATTCAATCGTGTCTTA 7682
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 ttaataatatccagcagagtgagaaagcgtttatcaattatgctcagttctctctt 567
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7683 APAGATTTTGTGTTCTCTATGGGATTCAATTTATTAACCTGATGTTATCATTTTCAGTT 7742
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 568 cctggcataatatgtatggtatttagcgcatctggtcttaggtcttttaactcggcgca 627
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7743 CCATGATTATCGGTATATTCTATGTGATATGGGGTTTGGGTTCTTAACAAAACAGCG 7802
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 628 caacaattgaatggttttctctctcaatgcgcgtcaaaagtatatgtgtctactgacg 687
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7803 CCTCAACTAAACGATATTCACATTACATGCCAGTAAAAAGTTTGATAGCAATCTTTATA 7862
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 688 ctccgatctcattcccttatgctcttcatt 717
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	7863	TTGCTATTAGTTATTCATGTTTTTCTGAT	7892
RESULT 11	AE005597	11760 bp	linear BCT 21-MAR-2001
LOCUS	Escherichia coli O157:H7 EDL933	genome, contig 3 of 3	section 216
DEFINITION	Escherichia coli O157:H7 EDL933		
ACCESSION	AE005597	AB005174	
VERSION	AE005597.1	GI:12518472	
KEYWORDS			
SOURCE	Escherichia coli O157:H7 EDL933.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
REFERENCE	1 (bases 1 to 11760)		
AUTHORS	Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grothbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., and Blattner, F.R.		
TITLE	Genome sequence of enterohemorrhagic Escherichia coli O157:H7		
JOURNAL	Nature 409 (6819), 529-533 (2001)		
MEDLINE	21074935		
PUBMED	11206551		
REFERENCE	2 (bases 1 to 11760)		
AUTHORS	Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grothbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., and Blattner, F.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA		
FEATURES	Location/Qualifiers		
source	1. .11760		
	/organism="Escherichia coli O157:H7 EDL933"		
	/strain="EDL933"		
	/serotype="O157:H7"		
	/db_xref="taxon:155864"		
	/note="enterohemorrhagic"		
	<1. .9623		
misc_feature	/note="O-island #148; Region of the EDL933 chromosome not homologous to E. coli K-12 MG1655"		
gene	129. .587		
	/gene="E5131"		
CDS	129. .587		
	/gene="E5131"		
	/function="orf; Unknown function"		
	/note="Residues 1 to 152 of 152 are 100.00 pct identical to residues 1 to 152 of 152 from GenPept 118 :		
	g1 3414913 gb AAC31524.1 (AF071034) L0045 [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="AAG58843.1"		
	/db_xref="GI:12518473"		
gene	/translation="MKKILSLIISNYCYASDCFEITGRAYNDPILIKAIWANESKCSGKSKTKNGTYDIGIMINSNHLKSFNEDDLNDACINISVAGYILASNIKSRGNTDVGAYNAGYFNPNAVELRQYAMKIYTKNKLKNEQIID"		
	complement(584. .1621)		
	/gene="escu"		
	/note="E5132"		
	complement(584. .1621)		
	/gene="escu"		
	/function="transport; Transport of large molecules; Protein, peptide secretion"		
	/note="Residues 1 to 345 of 345 are 100.00 pct identical to residues 1 to 345 of 345 from GenPept 118 :		
	g1 3414914 gb AAC31525.1 (AF071034) L0046 [Escherichia coli]"		
gene	/codon_start=1		
	/transl_table=11		
	/protein_id="AAG58844.1"		
	/db_xref="GI:12518474"		
gene	/translation="MSKTEKPTPKKLRDLKKKGDVTKSEVMAVOSLILSFSSFLYSGNSFIDVGLVNTTIDSLNRPFLYAIREILGAVINFLYILPISLIVFGVTTGYSQISFIVAEIKPSAQKISVKNLKNIFSVKISFELLKSVFKLVITLVIFYEMGHISANEPAFTGLNAYOALVVAFFVFLKMGVLFGYLLFSDFDFWFOHKGILKMKMSKDEVKREKDTGDNPEIKERRRLHSEIOSGSLANNIKKSTVIVKNPHTIALCLYKIGE TPLPLVETGDKAKALQIIKLAELYDIPVEDIPLARTLYKNHKQYITTEDFPFPAVAQLIRIAIDLDY"		
	complement(1614. .2390)		
	/gene="escr"		
	/note="E5133"		
	complement(1614. .2390)		
	/gene="escr"		
	/function="transport; Transport of large molecules; Protein, peptide secretion"		
	/note="Residues 1 to 258 of 258 are 100.00 pct identical to residues 1 to 258 of 258 from GenPept 118 :		
	g1 3414915 gb AAC31526.1 (AF071034) L0047 [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/product="escr"		
	/protein_id="AAG58845.1"		
	/db_xref="GI:12518475"		
	/translation="MNEIMTVSVSFYLCRLPCLPILPFTSCVLLSNFIRNSIMIAFTPLIVENTFSEKPSGIFQLTGALKEISGIFGLSFTILFWAIDAGQIITLAFSTISIFNPSIDSSITGVLYQFIVFVHGGIOSILDKLYSLYEILPLQANDTAFNALIDFLFSLWDSFKLMSVPMIIGFICDMGFGFLNKTAPQLNVTLSLPVKSLIAIFILLVHVFPDFITANIHSIDIDRSLPSMINE"		
	complement(2390. .2659)		
	/gene="escs"		
	/note="E5134"		
	complement(2390. .2659)		
	/gene="escs"		
	/function="transport; Transport of large molecules; Protein, peptide secretion"		
	/note="Residues 1 to 89 of 89 are 100.00 pct identical to residues 1 to 89 of 89 from GenPept 118 :		
	g1 2865277 gb AAC38370.1 (AF022236) Escs [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/product="escs"		
	/protein_id="AAG58846.1"	</	

Db	7863	TTGCTATTAGTTATTCATGTTTTTCTGAT	7892
RESULT 11	AE005597	11760 bp	DNA linear BCT 21-MAR-2001
LOCUS	Escherichia coli O157:H7 EDL933	genome, contig 3 of 3	section 216
DEFINITION	Escherichia coli O157:H7 EDL933		
ACCESSION	AE005597	AB005174	
VERSION	AE005597.1	GI:12518472	
KEYWORDS			
SOURCE	Escherichia coli O157:H7 EDL933.		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
REFERENCE	1 (bases 1 to 11760)		
AUTHORS	Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, S., K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., and Blattner, F.R.		
TITLE	Genome sequence of enterohemorrhagic Escherichia coli O157:H7		
JOURNAL	Nature 409 (6819), 529-533 (2001)		
MEDLINE	21074935		
PUBMED	11206551		
REFERENCE	2 (bases 1 to 11760)		
AUTHORS	Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, S., K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., and Blattner, F.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA		
FEATURES	Location/Qualifiers		
source	1. .11760		
	/organism="Escherichia coli O157:H7 EDL933"		
	/strain="EDL933"		
	/serotype="O157:H7"		
	/db_xref="taxon:155864"		
	/note="enterohemorrhagic"		
	<1. .9623		
misc_feature	/note="O-island #148; Region of the EDL933 chromosome not homologous to E. coli K-12 MG1655"		
gene	129. .587		
	/gene="E5131"		
CDS	129. .587		
	/gene="E5131"		
	/function="orf; Unknown function"		
	/note="Residues 1 to 152 of 152 are 100.00 pct identical to residues 1 to 152 of 152 from GenPept 118 :		
	g1 3414913 gb AAC31524.1 (AF071034) L0045 [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="AAG58843.1"		
	/db_xref="GI:12518473"		
	/translation="MKKILSLIISNYCYASDCFEITGRAYNDPILIKAIWANESKCSGKSKTKNGTYDIGIMINSNHLKSFNEDDLNDACINISVAGYILASNIKSRGNTDVGAYNAGYFNPNAVELRQYAMKIYTKNKLKNEQIID"		
	complement(584. .1621)		
gene	/gene="escu"		
	/note="E5132"		
	complement(584. .1621)		
CDS	/gene="escu"		
	/function="transport; Transport of large molecules; Protein, peptide secretion"		
	/note="Residues 1 to 345 of 345 are 100.00 pct identical to residues 1 to 345 of 345 from GenPept 118 :		
	g1 3414914 gb AAC31525.1 (AF071034) L0046 [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="AAG58844.1"		
	/db_xref="GI:12518474"		
	/translation="MSKTEKPTPKKLRDLKKKGDVTKSEVMAVOSLILSFSSFLYSGNSFDVGLVNTTIDSLNRPFLYAIKEILGAVLNIFLLYILPSLIVFGVTTGYSQISFIVAEIKPSAOKISVKNLKNIFSVKSFELLKSVFKLVITLVIFPMGHISANEPAFTGLNAYOALVVAFFVFLKMKGVLFGLFSDFWFQHEGILKKMKSKGEVKEAKDTGDNPEIKERRBLHSEIOSGLANNIKKSTVIVKNPHTIALCLYKIGE TPLPLVETGDKAKALQIIKLAELYDIPVEDIPLARTLYKNHKQYITTEDFPFPAVQLIRIAIDLDY"		
	complement(1614. .2390)		
gene	/gene="escr"		
	/note="E5133"		
	complement(1614. .2390)		
CDS	/gene="escr"		
	/function="transport; Transport of large molecules; Protein, peptide secretion"		
	/note="Residues 1 to 258 of 258 are 100.00 pct identical to residues 1 to 258 of 258 from GenPept 118 :		
	g1 3414915 gb AAC31526.1 (AF071034) L0047 [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/product="escr"		
	/protein_id="AAG58845.1"		
	/db_xref="GI:12518475"		
	/translation="MNEIMTVSVSFYLCRLPCLPILPFTSCVLLSNFIRNSIMIAFTPLIVENTFSEKPSGIFQLTGALKEISGIFGLSFTILFWAIDAGOIITLAFSTISIFNPSIDSSITGVLYQFIVFVHGGIOSILDKLYSLYEILPLQANDTAFNALIDFLFSLWDSFKLMSVPMIIGFICDMGFGFLNKTAPQLNVTLSLPVKSLIAIFILLVHVFPDFITANIHSIDIDRSLPSMINE"		
	complement(2390. .2659)		
gene	/gene="escs"		
	/note="E5134"		
	complement(2390. .2659)		
CDS	/gene="escs"		
	/function="transport; Transport of large molecules; Protein, peptide secretion"		
	/note="Residues 1 to 89 of 89 are 100.00 pct identical to residues 1 to 89 of 89 from GenPept 118 :		
	g1 2865277 gb AAC38370.1 (AF022236) Escs [Escherichia coli]"		
	/codon_start=1		
	/transl_table=11		
	/product="escs"		
	/protein_id="AAG58846.1"		

CDS		complement(3317..3931) /gene="Z5136" /function="orf; Unknown function" /note="Residues 1 to 204 of 204 are 99.50 pct identical to residues 28 to 231 of 231 from GenPept 118 : gil3414918[gb AAC31529.1] (AF071034) L0050 [Escherichia coli]" /codon_start=1 /transl_table=11 /protein_id="AAG58848.1" /db_xref="GI:12518478" /translation="MLTEDIPEKQLEIIQSDDIKLARKTYEHLRRARRSKSEILS TESRKKIARWKLWRRIKRNNKKLKDEYNQSVKWYDQLAEVLWMQIMMKIHLS LTNALSHSDYSLRNWDLLNEVRETLSNNIVGAITKNPDIKLDPGEANNIQILI NDANTPRKIIEENYEIRITIDPLEQISILLNSFKDNLYSLIQE" /gene="Z5137" /complement(3957..4556) /gene="Z5137" /note="Residues 1 to 199 of 199 are 100.00 pct identical to residues 1 to 199 of 199 from GenPept 118 : gil3414919[gb AAC31530.1] (AF071034) L0051 [Escherichia coli]" /codon_start=1 /transl_table=11 /protein_id="AAG58849.1" /db_xref="GI:12518479" /translation="MTIFNKIDNYKYKLINPIPMVHDEMGLDTGVNQVSFRRLRET SSRNOLNLKEIHDKIGSFLSDINKEGFYQSIGRTLIALLSSALFDIOCPDY I FSRLYLREITCYQNKLSPYWGQGCKPEYGGRFCAELIKYVGACGNLEWLFPADP PLWTIVKLLPKSGEIKPTHINDLFNLRLNKLPPVEL" /gene="Z5138" /complement(4553..4876) /gene="Z5138" /note="Residues 1 to 107 of 107 are 100.00 pct identical to residues 1 to 107 of 107 from GenPept 118 : gil2865273[gb AAC39366.1] (AF022236) Orf3 [Escherichia coli]" /codon_start=1 /transl_table=11 /protein_id="AAG58850.1" /db_xref="GI:12518480" /translation="MSIVSQTRNKELLDKRKIRSEIFAIAKKIIAEFDVKESVNELSEK AKTDPOAAEKLANLKBGYTGGERLYDSALSIEKLIETLSPARSKSQTWNQRNR NRKIV" /gene="Z5139" /complement(4880..5098) /gene="Z5139" /note="Residues 1 to 72 of 72 are 100.00 pct identical to residues 1 to 72 of 72 from GenPept 118 : gil2865272[gb AAC39365.1] (AF022236) Orf2 [Escherichia coli]" /codon_start=1 /transl_table=11
		Query Match 15.1%; Score 118; DB 1; Length 11760; Best Local Similarity 52.0%; Pred. No. 2,2e-18; Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;
Qy	208	tggttagggttagtcaccgagaggcgattatgtgtttttaattgggttggcgcgacg 267
Dd	2189	TTCACAGGAAATACCGCTAAAGAATAAAGTAATTGGTTTTCATTGGTTATCATTTACT 2130
Qy	268	gtcccttttgccggttgtatatgccgcccgttcctgcctgatcttaactgaatgcgcgaca 327
Dd	2129	ATTCTTTTTGGGCATAGATGCCGTGCACAGATPATTTGATCTACTAAGAGCGCTCAACA 2070
Oy	328	atgggttacgatataccaattctaacaatagaagtgcgaacctcaacttttggctgtgttttc 387

Db	2069	ATATCTTCAATTTTAAACCGTCCATAAGTATTCATCTTCTATCACTGGCGTTATCTTG	2010
QY	388	agccagtttttggtggttatattcttttaagcggcggaatggaggttatattataaacatt	447
Db	2009	TACCAATTTATCTCTGTGATCTTTGTTATTCATGCTGGGATACAAAGCATCTAGATAAG	1950
QY	448	ctgtatgagtcataatcaatattaccaccagggcgctactttattattattgacggcaatt	507
Db	1949	CTATATTATCTACGAGATATTACCATTCACAGCCGATATGCAATCAATCGTGTCTTA	1890
QY	508	ttaaatatataccaggcagagtgaggaaacgctttatcaattatgctcagttctctct	567
Db	1889	ATAGATTTTGTGTTTCTCTATGGATTCATTTATTAACATGATTTATCATTTTCAGTT	1830
QY	568	ctgcataatataatggttataggcagatcgtcttaggtcttttaaatcgctcgga	627
Db	1829	CCCATGATTATCGGTATATCTTATGTGATATGGGTTGGGTTTCTTTAAACAAACAGCA	1770
QY	628	caacaattgaatggtttttctctcaatgcgcgtcaaaagtatatgtgtctactgacg	687
Db	1769	CCTCAGCTAAACGATTTACATATTCACCTGCCAGTAAAGTTTGATAGCAATCTTTATA	1710
QY	688	ctctgctatcattcccttatgctcttcat	717
Db	1709	TTGCTATTAGTTATTCATGTGTTTTCCTGAT	1680

RESULT	12
AF071034/C	
LOCUS	AF071034
DEFINITION	Escherichia coli L0001 (yicJ) gene, partial cds; linear BCT 13-AUG-1998 complete sequence; CP4-like integrase (int), L0004, L0005, L0006, L0007, L0008, L0009, L0010, L0011, L0012, L0013, L0014, L0015, L0016, L0017, L0018 (escF), L0019, L0020 (espB), L0021 (espD), L0022 (espA), L0023 (sep), L0024 (escD), L0025 (eaeA), L0026 (orfU), L0027 (tfr), L0028, L0029, L0030, L0031 (sepQ), L0032, L0033, L0034 (escN), L0035 (escV), L0036, L0037 (sepZ), L0038, L0039 (escJ), L0040, L0041 (escC), L0042 (escD), L0043, L0044, L0045, L0046 (escU), L0047 (escT), L0048 (escS), L0049 (escR), L0050, L0051, L0052, L0053, L0054, L0055, and L0056 genes, complete cds; and L0057 (yicL) gene, partial cds.
ACCESSION	AF071034
VERSION	AF071034.1
KEYWORDS	GI:3414870
SOURCE	Escherichia coli.
ORGANISM	Escherichia coli
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
AUTHORS	1 (bases 1 to 45325) Perna, N.T., Mayhew, G.F., Posfal, G., Elliott, S., Donnenberg, M.S., Kaper, J.B. and Blattner, F.R.
TITLE	Molecular evolution of a pathogenicity island from enterohemorrhagic escherichia coli O157:H7
JOURNAL	Infect. Immun. 66 (8), 3810-3817 (1998)
MEDLINE	98339885
REFERENCE	2 (bases 1 to 45325)
AUTHORS	Perna, N.T., Mayhew, G.F. and Blattner, F.R.
TITLE	Direct Submission
JOURNAL	Submitted, 045 JUN-1998) Genetics Laboratory, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES	Location/Qualifiers
source	1..45325 /organism="Escherichia coli" /strain="EDL933" /strain="ATCC43895" /serotype="O157:H7" /db_xref="ATCC:43895" /db_xref="taxon:562" /note="Isolated from contaminated hamburger implicated in an outbreak of haemorrhagic colitis" complement(<1..557) /gene="yicJ"
gene	

CDS	complement(c1..557) /gene="yicJ" /codon_start=1 /transl_table=11 /product="L0001" /protein_id="AAC31536.1" /db_xref="GI:3414925" /translation="MYRMRYRSHQPSYKKEKQKSEVLSEKIGYGMGGAASHIIF DNVLMYMFYFDNFGIPAGFVGTMTFLVARALDAISDCMGLLADRTSRWKGFRPW LGGALPFGIVCVLSTPDLSNMGKMIYAATITLLTLYTVVNTIPYCALGVITNDP TORISIQSWRFVLATAGGMLSTVLP"
tRNA	723..902 /note="L0002" /product="tRNA-OTHER" 1083..2264 /gene="int" 1083..2264 /gene="int" /note="L0003" /codon_start=1 /transl_table=11 /product="CP4-like integrase" /protein_id="AAC31482.1" /db_xref="GI:3414871"
gene	1083..2264 /gene="int" 1083..2264 /gene="int" /note="L0003" /codon_start=1 /transl_table=11 /product="CP4-like integrase" /protein_id="AAC31482.1" /db_xref="GI:3414871"
CDS	2363..2713 /codon_start=1 /transl_table=11 /product="L0004" /protein_id="AAC31483.1" /db_xref="GI:3414872" /translation="MISSPOHTGDLMNKTRTTTPPRLECAQLIVDKGYSYQAS EAMNVGSTTLESWVRQLRRERQGIAPSTIPDQORIRELEKQVRRLEHNTILKKA TVALMSDSLNGSR" 2726..2908 /codon_start=1 /transl_table=11 /product="L0005" /protein_id="AAC31484.1" /db_xref="GI:3414873" /translation="MTATOLSASVLIMKYTAATVTGTENDAIRLRRESGCAKYAGR GTVGALGRARWLK" 2926..3579 /codon_start=1 /transl_table=11 /product="L0006" /protein_id="AAC31485.1" /db_xref="GI:3414874" /translation="MSRYRAGRLMKVNLSSCPGKHQVKNAROEHTCLPNLLERQFA VPEPRVWCCDTITYIAGNRWCYLAVVMDLPRFVIGWISLSNADTALISSALRMAE VRGPRVDFDSDGSGYTGKLYQQLLMRLYKQSVSRNGCWDSMPMERFRLKYE WYPTDGYTKQVARQOISSYILNYNSVRPHHYNGLTPTPESENRYHFYCKTVASIT" 3830..4204 /codon_start=1 /transl_table=11 /product="L0007" /protein_id="AAC31486.1" /db_xref="GI:3414875" /translation="MRTLDPDTHVREASRCPSPITIWOTLLGRLLDOHYGLTNDTPFA DERVTEOHTAEGISLCAVNFELVEXYVLTVPDQPGFSACTRSQOLINSIDILARRATG LATRDNYRFTVNNITLGKHPK"
CDS	4201..4692 /codon_start=1 /transl_table=11 /product="L0008" /protein_id="AAC31487.1"
CDS	4704..4901 /codon_start=1 /transl_table=11 /product="L0009" /protein_id="AAC31488.1" /db_xref="GI:3414877" /translation="MTSLTPEALDILIAWLQDNIDSESGIIFONDEDKTDSSAALLPC IFQVREDVFTLRQLQLQNR" 4986..5327 /codon_start=1 /transl_table=11 /product="L0010" /protein_id="AAC31489.1" /db_xref="GI:3414878" /translation="MSGIISRPEVDTHGTVICSTSIIRHVTVVRNALQOETELIRQ LAEISVLTAADIGGKTALDRAMKQDPHCGCWLMEKPEATAMKAITRNLDCETWRDLQMR SGMLSMDAOG" complement(5324..5554) /codon_start=1 /transl_table=11 /product="L0011" /protein_id="AAC31490.1" /db_xref="GI:3414879" /translation="MKSPSPVSHQVVDNFTABFTGGIGFVIPPFGPLKTLITPRSNTS SLLMCSCLNAFMIFASLISGILSYSSERYHVS" 5528..5713 /codon_start=1 /transl_table=11 /product="L0012" /protein_id="AAC31491.1" /db_xref="GI:3414880" /translation="MWDWRFRHLISGMQADRLADLERMLHLLSGRPLNNRGNITIN LDHLSQSVGGRYEDK" 6083..6484 /codon_start=1 /transl_table=11 /product="L0013" /protein_id="AAC31492.1" /db_xref="GI:3414881" /translation="MCTKVSDMKNVTPGRRGCGPNYPPEFKOQLVAACEPGJISLK LALENGINANLLFKWQQRWEGKLLIPSSSPQLLPVTLDAAAEQPSLAEDPETLSI SCVETPRHGTLRFGNVSEKLLTLTIQELKR" 6481..6828 /codon_start=1 /transl_table=11 /product="L0014" /protein_id="AAC31493.1" /db_xref="GI:3414882" /translation="MIPLPSTGKIWLVAGITDMRNGFNGLAAKVQTALKDDPMGSHVF IFRGRSGSOVKLLMSTGDCGLLTTLKRLERGFAPWSARDGKVFELTQALNMLLEGIDW RQPKLLTISLTM" 6878..8416 /codon_start=1 /transl_table=11 /product="L0015" /protein_id="AAC31494.1" /db_xref="GI:3414883" /translation="MNDISDDIFLLKQRLAQEALIHAEQKLSNRREIHDHQAOL DKLRRNMFGRSEKSVRRIAQMPADLNRIKQESDTLGRVYPAVQRPILRQTRKPPF PESLPDERKRLPAAPCCPNCGSLSYAGEDAEQLLEMRSAFRLRTVREKHACTQC DAIVQAPASRPRIERGIAGPGLLARVLTSKYAEHTPLYRQSEIYRGQVELRRSLLSG WYDACCRLSPLEELHGYVMTDGLHDDTPQVLLPGNKTKTGRWAYVRDRNA GSALAPAVFAYSPDRKGHPQTHLACFSGLQADAYAGNELYRGGITAEACWAHA RRKIHDVHVRIPSALEAEQIGQLYAEADIRGMPARQRAERQRTKPLLSLES WLREKNKTLRSSELAFAFAYALNQPALTYANDGWVEIDNNIAFNALRAVSLSKHN FLFFGSDHGGERGALLYSLIGTCKLNDVDPESYLRHVGLVIADWPVNVYSSELLPWRIA LPAE"
CDS	complement(9281..10027)

```
/codon_start=1
/transl_table=11
/product="L0016"
/protein_id="AAC31495.1"
/db_xref="GI:3414884"
/translation="MLNGISNAASTLGRQLVGIASSVAGCTGFSVAPOAVRLTPVK
VHSPFSGSSNNVARTIFNVSVQVTFPSRPAPPPTSGQASGRPLPPIAQLAKE
HLAAVEKSGPALGFKPARQAPPPTSGQASGRPLPPIAQLAKEHLAAVEKSGP
EALGFKPARQAPPPTSGQASGRPLPPIAQLAKEHLAAVEKSGPALGFKPARQA
PPPTGPGSLPPLAQLKDHDLAAVEQSKG"
complement(10112..10390)

CDS
Query Match      15.1%; Score 118; DB 1; Length 45325;
Best Local Similarity 52.0%; Pred. No. 2e-18;
Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 208 tggtaggttagtcaccggagggattggttttttaattggtttggtgctgagcg 267
DB 36886 TTGACAGGAATAGCGCTAAAGAAATAAGTATTGGTTTTCATTTGGGTATCATTTACT 36827
QY 268 gtcccttttggccgtgatgacgggggtttctgcttgatcattacgtagcgcgaca 327
DB 36826 ATTCCTTTTGGCAATAGATCGGCTGGACAGATATTGATACTCTAGAGGCTCAACA 36767
QY 328 atgggtacgatatacaattctcaatagagctgaaacccctcatttttggctgttttc 387
DB 36766 ATATCTTCAATTTTAAACCGGCCATAGTATTGATCTCTTATCATTGCTGCGCTTATCTTG 36707
QY 388 agccagtttttgggtatttcttataagcggcgatggagtttatataaacatt 447
DB 36706 TACCAATTTATCTGTGTATCTTTGTTTATTCATGCTGGGATACAAAGCATCTAGATRAAG 36647
QY 448 ctgtatgagtcatacaattattaccacagggcgatggttttataataaacatt 507
DB 36646 CTATATTATCTACAGATATTACCATACAGCCGATATTGCAATCAATGCTGCTTTA 36587
QY 508 ttaaatatataccagcagagtgaggaacgtttatcaattatgctcagtttctctctt 567
DB 36586 ATAGATTTTGTCTCTATGAGGATCATTTATTTAAACGTGATTTATCATTTTCAGTT 36527
QY 568 ctctgcaataatgatgattagccagatcgtgcttttagtcttttaaatcggtcgaca 627
DB 36526 CCATGATATTACGGTATATCTTTATGATATGTTGGGTTTCTTTAAACAAACAGCA 36467
QY 628 caacaatgaatgtgttttcttccatgcgctcaaaagtatatgttcttactgacg 687
DB 36466 CCTCAGCTAAAGCTATTACATATTACATGCCAGTAAAGTTTGATAGCAATCTTTATA 36407
QY 688 ctctgctatcattcccttctgcttctcat 717
DB 36406 TTGCTATTAGTATTATCATGTTTTCCTCGAT 36377

RESULT 13
LOCUS AP002566 280900 bp DNA linear BCT 07-MAR-2001
DEFINITION Escherichia coli O157:H7 DNA, complete genome, section 17/20.
ACCESSION AP002566 BA000007
VERSION AP002566.1 GI:13363930
KEYWORDS
SOURCE
  Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RIMD 0509952)
  DNA.
ORGANISM
  Escherichia coli O157:H7
  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
  Escherichia.
REFERENCE
  1 (sites)
  Kurokawa,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
  Yamamoto,C.H., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
  Sasakawa,C., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
  Complete nucleotide sequence of the prophage VT2-Sakai carrying the
  verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
  derived from the Sakai outbreak

TITLE
```

JOURNAL
MEDLINE
REFERENCE
AUTHORS

Genes Genet. Syst. 74 (5), 227-239 (1999)

2 (sites)

Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M.,
Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and
Hayashi,T.

TITLE

Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655

Syst. Appl. Microbiol. 23 (3), 315-324 (2000)

20557356

JOURNAL
MEDLINE
REFERENCE
AUTHORS

3 (sites)

Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S.,
Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T.,
Yamamoto,C.H., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and
Shinagawa,H.

TITLE

Complete nucleotide sequence of the prophage VT1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
O157:H7 strain derived from the Sakai outbreak

Gene 258 (1-2), 127-139 (2000)

20564182

JOURNAL
MEDLINE
REFERENCE
AUTHORS

4 (sites)

Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K.,
Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T.,
Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C.,
Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and
Shinagawa,H.

TITLE

Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12

DNA Res. 8 (1), 11-22 (2001)

21156231

JOURNAL
MEDLINE
REFERENCE
AUTHORS

5 (bases 1 to 280900)

Ohnishi,M., Kurokawa,K., Makino,K., Yasunaga,T., Shinagawa,H. and
Hayashi,T.

TITLE

Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail:ken@gen-info.osaka-u.ac.jp,
URL:http://www.gen-info.osaka-u.ac.jp/; Tel:81-6-6879-8365,
Fax:81-6-6879-2047)

genome project.

COMMENT
FEATURES
source

Location/Qualifiers
1..280900
/organism="Escherichia coli O157:H7"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
complement(143..616)
/gene="ECs4456"
complement(143..616)
/gene="ECs4456"
/note="similar to YIAI_ECOLI YIAI_ECOLI q112367245 percent
identity 98 in 157 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="BAB37879.1"
/db_xref="GI:13363931"

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS


```

/product="putative transcriptional regulator"
/db_xref="GI:13363932"
/translation="MLELSMTLPKIQVNGGLFISRGVGRHSARRLESWEIIFVEKGRLL
TTOEERIFLVDAESGLLLPNRQHVGVDFPADLKFWLWHEVKAPEADNPRLWTLHLS
VPOQTKVDPQALVDFQFLSEQKHQRTPALEFIVLLILOQLFRDLRDEQDAGAG
TTLAKAOKLIEQTVHRPLSTALLARELYCNADYLGRRYRVFHLTLTAIRQVRRE
AKLLISEARSUTEVAALCGFNDVGIFRQIRKHTGLTTPAAWKRRYSKEHINS"
1901..3301
/gene="ECs4458"
1901..3301
/gene="ECs4458"
1901..3301
/notes="probable permease similar to hypothetical proteins
[Salmonella typhimurium] gi|7442781|p|l|c65167 percent
identity 37 in 444 aa, also similar to transport proteins
(putative symporters) e.g. YicJ [Escherichia coli (K-12)]
gi|2851421|sp|p31435|YICJ_EC"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="putative permease"
/protein_id="BAB37881.1"
/db_xref="GI:13363933"
/translation="MTSTPTTADIAKIKIDDALSVREKIGYGLGDAGGTIVTCLIMN
ETFEVTVDFGTPALVGLFELALRVDFDALSDPMGVADIQTSRGRPRPQLHIAI
PGIIGILFTVPDASGVKIAWAGTGYLLSVGTAINVPCARINTTRHNEVIS
COSMRVLCGVAGFIVSGVLPWMAVFCGNAARGVQGVLCALAVNMFCCFFFW
RRVPVLSMTKRTLRHLAGLRNNQDLMVLFLLINVFNIIRGGGYMFITYVLQG
STGYTSLFTMTVFASIGSVIVSPLTRRFDVKIYYTNLLAALAVLWMLPSPGA
YOTLMLAVILGNVGLGTLPLHFSLMAFADDDYGEWKTVHVRSSGMNFAPNLFKILAW
ASSAGTISLFIYVAYQGVENQTSASSLGGIAAMETLLPALPHLLLAGAIRCKLNP
MMSRIATDQRHVQP"
3312..5282
/gene="ECs4459"
3312..5282
/gene="ECs4459"
/notes="unknown similar to hypothetical protein [Thermotoga
maritima] gi|7452109|p|l|F7395 percent identity 37 in
635 aa"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="BAB37882.1"
/db_xref="GI:13363934"
/translation="MNISEVDLKLTVSDPFLGQVQOOLVRDVVIVPQMDALNDRIPEA
EPSHAENFRIAAGLQEGEFGYGVFQSDVAKWLEAVAMSLQKPDALFKTADIEVIE
LIASAQEDGYLNTYTVKAPPEERWSNLAECHELICAGHLIEAGVAFQATGRRLLG
VYCRADHIDSVEGPDSEKILGYPGHPETEALMRLEYTEERPYLALTNYVEQORGA
OPHYDOEYKRGQTSWHITYGPANVVKDKAYSQAHPLAQOQTALGHAVRFYLMTG
VAHLARLHDDSRKQDCLRLMNNMAORQLYITGGIGSQSSGEAFSPDLPNDTVYAE
SCASIGLMFARMLMEGDSQYADVMERALYNTVIGMALDGKHFYFNPLEVHPKS
LKFNIHYDKPTIRQWFGCACPPNIARVLTSIGHLYTPREDALYINITYAGNSMEV
PVENGTLRHVSQNYPWQSOVTVAVESQPVHRITLALRPLDWCTQPOITLNGVEQDN
IRKGYLHTRREQEGDTLNTLPMVPRVRYGNPLRVHAGKVALORGLVYCLEKXADN
GESLHNLMLPTDAPFTTFEGKGLFSHKILLQAPGYRYEQSNPQOPLMHWDSAPAKRQ
TOTLTFIPFWSANRGEGRWVNEEKHCP"
complement(5412..6152)
/gene="ECs4460"
complement(5412..6152)
/gene="ECs4460"
/notes="similar to YIAT_ECOLI gi|1790010 percent identity
99 in 246 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="putative outer membrane protein"
/protein_id="BAB37883.1"
/db_xref="GI:13363935"
/translation="MLINRNIVALFALPFWASATSELSIGAGAAYNESPYRGYNENT
KAIPLSGVADWYVHERWGTVKASAAADVLQNSNGWVGLSVFHKMIDIGRLSTPALGV
RSTAMAGVAVYHHERWGTVKASAAADVLQNSNGWVGLSVFHKMIDIGRLSTPALGV
LYDENFSDDYVIGISESRSSGLASYSAQDAWVPVYSLTAKYPIGEHVVLMSAGYS

```

```

ELPEETDSPMDIRNESFTFTGVSWRP"
6276..7250
/gene="ECs4461"
6276..7250
/gene="ECs4461"
/notes="similar to YIAU_ECOLI gi|1790011 percent identity
99 in 324 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="putative transcriptional regulator LYSR-type"
/protein_id="BAB37884.1"
/db_xref="GI:13363936"
/translation="WTKLQKRYRELKIIIVIASENISUAATVVLGTAQNAVSKYLADF
ESKVLKRVFDRTRQMLTLFGTALLPYINDLDRNEQLNNFIADYKHKRGKRVTIYA
PTGIIITLTKHVIDIKIDCDITSLKTNLERNAFEGVEFFDDCDVLSVAPPKDE
SVLASPTQAVTAVASQRYLEKHPISRPDELHSHCILLDSMMIDIDANLWRPNVAGS
KEVRDYRVKGVYCDNTQSALEARNHLGTVFAPDKSVOSLDQDGLVPCFQHPYEMW
LDLVAIFRKREYQVRYVLDENLREIRIQLAQSOQLRPEQNAESD"
complement(7247..8383)
/gene="ECs4462"
complement(7247..8383)
/gene="ECs4462"
/notes="similar to YIAV_ECOLI gi|1790012 percent identity
99 in 378 aa, similar to YIBH_ECOLI gi|1790024 percent
identity 76 in 378 aa (Conserved in E.coli K-12)"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="putative membrane protein"
/protein_id="BAB37885.1"
/db_xref="GI:13363937"
/translation="MDLLIILTVYVAFAMFKIFKIPVKNWTTIPTAALGGIFIVSGLI
LLMNYNHPYTFKAQKAVISIPVQTVGVIEVTDKKNTLTKKGEVLFRIDPTRYQAR
VDRLMADIVTAEHKORALCAELDEMAANTKAKATROKFAKQYRVARGSQAKVNPFS
ERDIDVARNQLAQEASVKSAAEQIKQISQLDSLVLGEHSQIASLKNQLAEAKYNLE

```

Query Match 15.1%; Score 118; DB 1; Length 280900;
Best Local Similarity 52.0%; Pred. No. 1.7e-18;
Matches 265; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

Qy 208 tggtagggtagtcaccggagggtaggtattatgggttttttttaattgggttttgcggcg 267
Db 129813 TTGACAGAAATAGCGCTAAAGAANAATAGTATTGGTTTTTCATGGGTATCATTTACT 129754

Qy 268 gtcccttttgggcgttgatagtcgggggtttctgtctgatacttaccgtggcgcgaca 327
Db 129753 ATTCTTTTTTGGGCAATAGATGCGCTGGACAGATTATTGATCTCTTAAGAGGCTCAACA 129694

Qy 328 atggggtacgatattcaatttcataatagaagcgtgaacccctcaacttttggcttgcctttc 387
Db 129693 ATATCTTTCAATTTTAAACCGCTCCATTAAGTCATTCTTCATCACTGGCGTTATCTTG 129634

Qy 388 agccagttttgtgtttattttcttataaagcggcgagcctggaggtttatataaacatt 447
Db 129633 TACCAATTTATCTCTGTGATCTTTGTTTATTCATGTGGGATACAAAGCATCTAGATAG 129574

Qy 448 ctgtatgagtcatacaatatttaccacccggcgctctttattatttggacggcgcaattt 507
Db 129573 CTATATTTATCTTACGAGATATTACCATTACAGCCGATATTGCAATTCGTCGCTTTA 129514

Qy 508 ttaaaatataccagggcagagtgaggaacgctttatcaaatatgltgcagtttctctctt 567
Db 129513 ATAGATTTTGTGTTTCTCTATGGGATTCATTTATTAACATGATGTTATCATTTTTCAGTT 129454

Qy 568 cctggccataatgtaggtattagccgactctggcttttaggtcttttaaacctcgcgca 627
Db 129453 CCCATGATTTATCGGTATTTATCTTATGTGATATGGGTTTGGGTTTCTTAACAAACAGCA 129394

Qy 628 caacaattgaatgtgttttcttctcaatgcgcgtcacaagaatgatatgttctactgacg 687
Db 129393 CCTCAGCTAANCGTATTACACATTATCACTGCCACGTAAAGATTGATAGCATCTTTTATA 129334

QY	688	ctctgacatccatcccttctgcttctcat 717
Db	129333	TTGCTATTAGTATTATCATGTTTCTCTGAT 129304
RESULT	14	
AF200363		
LOCUS		
DEFINITION		
Escherichia coli Ler (ler), EscR (escR), EscS (escS), EscT (escT), EscU (escU), Cesd (cesD), EscC (escC), EscJ (escJ), SepZ (sepZ), EscV (escV), EscN (escN), SepQ (sepQ), Tir (tir), Cest (cest), Intimin (eae), EscD (escD), SepL (sepL), EspA (espA), EspD (espD), EspB (espB), EscF (escF), EspF (espF), and putative transposase genes, complete cds; Lifa (lifa) gene, partial cds; and unknown genes		
AF200363		
AF200363.1	GI:13447694	
Escherichia coli.		
Escherichia coli.		
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia		
1 (bases 23013 to 24048)		
Agin,T.S., Canteley,J.R., Boedeker,E.C. and Wolf,M.K.		
Characterization of the eaeA gene from rabbit enteropathogenic Escherichia coli strain RDEC-1 and comparison to other eaeA genes from bacteria that cause attaching-effacing lesions		
FEMS Microbiol. Lett. 144 (2-3), 249-258 (1996)		
97055784		
8900070		
2 (bases 1 to 37889)		
Elllott,S.J., Wainwright,L.A., McDaniel,T.K., Jarvis,K.G., Deng,Y.K., Lai,L.C., McNamara,B.P., Donnenberg,M.S. and Kaper,J.B.		
The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic Escherichia coli E2348/69		
Mol. Microbiol. 28 (1), 1-4 (1998)		
98254123		
3 (bases 1 to 37889)		
Zhu,C., Agin,T.S., Elliott,S.J., Johnson,L.A., Thate,T.E., Kaper,J.B. and Boedeker,E.C.		
Complete Nucleotide Sequence and Analysis of the Locus of Enterocyte Effacement from Rabbit Diarrheagenic Escherichia coli RDEC-1		
Infect. Immun. 69 (4), 2107-2115 (2001)		
21153569		
11254564		
4 (bases 1 to 37889)		
Boedeker,E.C., Zhu,C., Elliott,S.J., Tonia,T.S., Johnson,L.A., Thate,T.E. and Kaper,J.B.		
Direct Submission		
Submitted (01-NOV-1999) University of Maryland School of Medicine, Center for Vaccine Development, 685 West Baltimore St., Baltimore, MD 21201, USA		
Location/Qualifiers		
1..37889		
/organism="Escherichia coli"		
/strain="RDEC-1"		
/serotype="O15:H-"		
/db_xref="taxon:562"		
/clone="pDK5"		
/note="Rabbit EPEC, REPEC"		
complement(<1..396)		
/organism="Escherichia coli"		
/strain="RDEC-1"		
/db_xref="taxon:562"		
/clone="pDK5"		
/insertion_seq="IS2"		
1032..35232		
/note="RDEC-1 locus of enterocyte effacement (LEE) pathogenicity island core region; similar to sequences presented in Genbank Accession Numbers AF022236 and AF071034"		
complement(1209..2027)		
misc_feature		
source		
source		
misc_feature		
gene		
CDS		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118</		

```
/gene="escr"
/codon_start=1
/transl_table=11
/product="Escr"
/protein_id="AAK26700.1"
/db_xref="GI:13447701"
/translation="MSQMTIGSOPIFLIIVFLLSLLDIFVVGTSFLKISIVIGIL
KNAIGQQVPPNALTVSLLILTMFIMSPIILQINDNISQEPINYPDSDFFOKVDEKI
LSPYRGFLKTEKENVEFERAAOKKLGNETILKDSLFILLPAFTMGOLEAERKIG
FLYLPFIAIDIIISNILLALGNMVMSPVTISIPRKILLFILVGGWQKLFEFLVYN"
7351..7620
/gene="escs"
7351..7620
/gene="escs"
7351..7620
/codon_start=1
/transl_table=11
/product="Escs"
/protein_id="AAK26701.1"
/db_xref="GI:13447702"
/translation="METGIFVQLCVQFLMIIFILSLPTVIAASVGVIIISLVOAITQL
QDOTLPFLKIIIAVFATLALTYHMGTTIINFSSIIIFEMIPKVG"
7620..8396
/gene="escT"
7620..8396
/gene="escT"
/codon_start=1
/transl_table=11
/product="EscT"
/protein_id="AAK26702.1"
/db_xref="GI:13447703"
/translation="MNEIMTVIVSSPFCILRPLCMFIILPIESTGVLLSNFRNSIMI
ATFLPIVENYFSEKPSGIFQITGVILKEISIGFIFGLSFILFMAIDAQOIIDT
LRGSTISFNPSIDSSGITGVILYOFITVIFGVHGGQSVDLKILYLSYEILPLQAD
IAFNALVDFLSLMDSFILKMSFIPMIIGIIFLDMGFGFLNKTAPOLINVTLSLP
VKSILAFILLIVHFVPDFITANIHSIDIIRSLPSMINE"
8389..9426
/gene="escU"
8389..9426
/gene="escU"
/codon_start=1
/transl_table=11
/product="EscU"
/protein_id="AAK26703.1"
/db_xref="GI:13447704"
/translation="MSEKTEPTPKLRLKKGDVTKSEVMAAVOSLILFFSFLSY
GASFVEVVELVNTTIDSLNRPFLYAIREILGAVLNIFLLYILPILSVFVGTVTGV
SOTGFLEVEKIKPSAQKISVKNLNKINFSIKSIFELLSKSVFKLITLIFYPMGHSY
ANEFNFTRLNAVOTLVVAVFVFLWKGVLFGYLSVDFDFQKHEGLKMKMSKO
EYKREAKTDGNEIKGERRLHSEIQSGSLANNIKSVIVKNPHTIAICLYKIGE
TLPVLETGKAKALQIIKLAELDIPVIEDIPLARSLYKNIHKQYTEDFEPVA
QLIRIADLDY"
complement(9423..9881)
/notes="r0f3"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK26704.1"
/db_xref="GI:13447705"
/translation="MKKIIISILISNSCYASDCFEITKAYNIDPILKAIAMNESK
NKNKISKINKNTYDGIQIMQINSHLDLISKENISDOLLNDACINISVAGYILASN
IKSRGNTDAVGAYNAGYFNPNAVELRQYANKIYKTYTKLNNEQID"

Query Match 14.6%; Score 113.8; DB 1; Length 37889;
Best Local Similarity 52.7%; Pred. No. 2.1e-17;
Matches 247; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

Qy 208 tgggttaggttagtcacggagaggtgattattggttttttaattggtttttgtgcggcg 267
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7821 TTGACAGGCATAGTGCATAAGTAAGTATAGGCTTTTCATGGGTATATCTACT 7880
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 268 gtcccttttggcgctgatagcggtgttctgcttctgatactttacgtgcgcgcaca 327
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7881 ATTCTTTTGGGCAATAGATGCGGCTGGCAGAGATTATGCATCTTTAAGAGGTTCAACA 7940
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Qy 328 atgggtacgatattcaattctacaatagaagctgaacccctcaactttttggttgcttttc 387
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 7941 ATATCTTCAATTTTAAACCGCTCTATAAGTAGATTATCATCTTCTATCACTGGGTATTATTTA 8000
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 388 agccagtttttggtgttattttctttataaagcgcgagctggagtttataataaacatt 447
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8001 TACCAATTTATCACTGTGATCTTTGTTTATTTCATGGTGGGATACAAACGCTTCTCGATAAG 8060
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 448 ctgtatgaglcatacaatattaccaccagggcgctactttattatttaccgcgcaattt 507
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8061 TTGTATTTTCCACGAGATATTACCATTTACAAGCCCATATTCGATTCATTCGTCGCTTTA 8120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 508 ttaaaatataccagggcagagtgaggaagcgttttatcaaatatgtgcagttttctctt 567
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8121 GTAGATTTTATTTTCGCTATGGGATTCATTTATTAAGCTGATGTATCATTTTCAATT 8180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 568 cctggcctaataatgtatgtattagcgcgtatcgtctttaggtcttttanaatcgtgcgca 627
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8181 CCCATGATTTATCGGTATCTCTTATGTGATGGGTTTGGGTTTCTCAATATAAACACGC 8240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 628 caacaattgaatgtgtttttctctcaatgcgcgtcacaagatatattgg 676
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8241 CCTCAGCTAAACGTATTACATTTATCACTCCCGCTAAAAAGTTTGATAG 8289
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 15

```
ECO277443
LOCUS
DEFINITION
60073 bp DNA for locus of enterocyte effacement II (LEE
Escherichia coli DNA linear BCT 05-DEC-2001
II).
```

AJ277443

AJ277443.1 GI:17384625

VERSION

KEYWORDS

cadC gene; CadC protein; cesD gene; CesD protein; cest gene; Cest protein; eae gene; eaeA gene; eaeB gene; eaeC gene; eafA-lifa-tox gene; Efa1-Lifa-Tox protein; ent gene; Ent protein; escC gene; EscC protein; escJ gene; EscJ protein; escN gene; EscN protein; escP gene; EscP protein; escS gene; EscS protein; escT gene; EscT protein; escU gene; EscU protein; escV gene; EscV protein; espA gene; EspA protein; espB gene; EspB protein; espD gene; EspD protein; espE gene; EspE protein; espF gene; EspF protein; int gene; Integrase; intimin; ler gene; Ler protein; orfU gene; P4 gene; sepA gene; sepB gene; sepC gene; sepD gene; sepE gene; sepF gene; sepG gene; sepH gene; sepI gene; sepL gene; sepM protein; sepO gene; SepO protein; sep2 gene; Sep2 protein; st01 gene; ST01 protein; st02 gene; ST02 protein; st03 gene; ST03 protein; st04 gene; ST04 protein; st06 gene; ST06 protein; st07 gene; ST07 protein; st08 gene; ST08 protein; st13 gene; ST13 protein; st14 gene; ST14 protein; st15 gene; ST15 protein; st18 gene; ST18 protein; st20 gene; ST20 protein; st22 gene; ST22 protein; st25 gene; ST25 protein; st26 gene; ST26 protein; st28 gene; ST28 protein; st29 gene; ST29 protein; st30 gene; ST30 protein; st39 gene; ST39 protein; st41 gene; ST41 protein; st43 gene; ST43 protein; st44 gene; ST44 protein; st46 gene; ST46 protein; st47 gene; ST47 protein; st48 gene; ST48 protein; st50 gene; ST50 protein; st51 gene; ST51 protein; st52 gene; ST52 protein; st53 gene; ST53 protein; st54 gene; ST54 protein; st55 gene; ST55 protein; st56 gene; ST56 protein; st57 gene; ST57 protein; tir gene; transposase. Escherichia coli.

SOURCE

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Benkel, P. and Chakraborty, T.
Genetic organisation and sequence of the LEE II locus in Shiga toxin-producing Escherichia coli
Unpublished
2 (bases 1 to 60073)
Benkel, P.
Direct Submission
Submitted (19-APR-2000) Benkel P., Justus-Liebig-Universitaet
Gießen, Institut fuer Medizinische Mikrobiologie, Frankfurt

Strasse 107, Giessen, 35392, GERMANY
related accession number AE000486, Escherichia coli K12 genome.
COMMENT: Location/Qualifiers
FEATURES
source
1. 60073
/organism="Escherichia coli"
/strain="413/89-1"
/serotype="O26:H"
/isolate="cali"
/db_xref="taxon:562"
/country="Germany"
/note="Shiga toxin-producing Escherichia coli (STEC)"
complement(1..49)
/gene="cadC"
/gene="cadC"
complement(41..49)
/gene="cadC"
/function="regulator of amino acid degradation"
/codon_start=1
/transl_table=11
/product="CadC protein"
/protein_id="CAC81838.1"
/db_xref="GI:17384626"
/translation="MQPVRVGEWLVTPS"
complement(552..687)
/rpt_type=DIRECT
complement(666..>687)
/gene="pheU"
/product="transfer RNA-Phe"
complement(666..687)
/gene="pheU"
complement(710..1000)
/gene="st01"
complement(710..1000)
/gene="st01"
/codon_start=1
/transl_table=11
/product="St01 protein"
/protein_id="CAC81839.1"
/db_xref="GI:17384627"
/translation="MDGRSRSDTDDVLLRISLRGCRFTSTANRELSLTTWRKRI
RQHCPVLNRPKPGCTFLCSPDPCDSLSISGKTAGLIEGAFLLRPNLH"
complement(1072..1437)
/gene="st02"
complement(1072..1437)
/gene="st02"
complement(1702..1437)
/gene="st02"
/codon_start=1
/transl_table=11
/product="St02 protein"
/protein_id="CAC81840.1"
/db_xref="GI:17384628"
/translation="MIDVLGPEKRRRTTQEKIATVQSPFGMTVSLVARQHVAA
QLFWRKQYQEGSLTAVAGEQVVPASELAAMQIKELQRLCKKTMEALLKEAVE
YGRKKWIAHAPLLPGDGE"
complement(2311..3129)
/gene="st03"
complement(2311..3129)
/gene="st03"
complement(2311..3129)
/function="putative outer membrane protein"
/codon_start=1
/transl_table=11
/product="St03 protein"
/protein_id="CAC81841.1"
/db_xref="GI:17384629"
/translation="MRRLFIAPFFIAMFVTPIALILFVFDYGYKKSQBYSLFKQL
PNVNLSSLNKEDFLYATONSQAQLKIAISGDIIGRAPLFTISAETHEIQGN
IFRAVDEKYSLEFFFTVKDMHLSLNKIQLEKFNKNRPFGLAWAGRMFLVAK
NLTGEMFAYQLSDPLLSAKQVITPELNDIHVKDLSGLDFNNEISLWLSDESKLKF
NLTGEMFAYQLSDPLLSAKQVITPELNDIHVKDLSGLDFNNEISLWLSDESKLKF
complement(3276..4472)
/gene="st04"
complement(3276..4472)
/gene="st04"
complement(3276..4472)
/function="Involved in bacterial invasion"
/codon_start=1

repeat_region
misc_RNA
gene
CDS
/transl_table=11
/product="St04 protein"
/protein_id="CAC81842.1"
/db_xref="GI:17384630"
/translation="MISVVFIIINETGESLMINGLNNSASLVDAIRINSDFKKQ
WDMSCAKLLKLVLSFGWNETYRSRQTELLTVLEPVSAPNELGIYANFADG
SSLRVSNSELVAEINTPDNEKILMLLESNEQRLLQSLPINLHMPYIQVHRALSK
MDLTDHSMHNLSTFKLSATLPHINTQTDPLSGPTPFSSMFMDFRGNGAKLSLN
GVDPDAQKFRDAGLKDTHSSILARNVINGISRHHAKQIAPRESSGSKQKAEVVE
FLCHPEAATAICSAFYOSFNPALMLTHTRIQSAREYNVERSLEDVFNACINISISQSP
DGSITHVASTGILLIMAPEDRNELGMLTNTTSYEVPOGVKCEIDEMVRTILOPRYGASE
TYLKN1"
5612..6001
/gene="ler"
5612..6001
/gene="ler"
/function="putative transcription activator"
/codon_start=1
/transl_table=11
/product="ler protein"
/protein_id="CAC81843.1"
/db_xref="GI:17384631"
/translation="MRRLFIMNMTSHTTSPYIQLIEQLVLOQAKRLRGEIIOV
IESIQKQIYYNITLQELGYTNVDDGLARNSSKGVYRNNEGQTSWGVGRPWIK
EALLNGMKKEDELVKDTEDEIIPLKN1"
6016..6234
/gene="st06"
6016..6234
/gene="st06"
/function="putative secreted protein"
/codon_start=1
/transl_table=11
/product="St06 protein"
/protein_id="CAC81844.1"
/db_xref="GI:17384632"
/translation="MITITELEDEIKKNEAANIFTEKINDKKNETHKMKHPLDKVT
YNKAKELLACDAAIRIIEIMLIRINN"
6558..7157
/gene="st07"
6558..7157
/gene="st07"
/function="unknown"
/codon_start=1
/transl_table=11
/product="St07 protein"
/protein_id="CAC81845.1"
/db_xref="GI:17384633"
/translation="MTILNKIDYNYKLIYPIIMVHDEWLGLTGVNQVSRHILRES
SSTRQNLKILRQETQDKISGVELSDINKEGFLYQSIGKIRLLASSALFEIQPDYI
FSRLYELIIRIIGYONVKQLSFYWGQCKPEYGEERFCSELIKYAGNLEWLFSDN
PLMTIVKLLPKSGBIKPTINDLFLNRLNKLILPYETL"
7102..7797
/gene="st08"
7102..7797
/gene="st08"
/function="unknown"
/codon_start=1
/transl_table=11
/product="St08 protein"
/protein_id="CAC81846.1"
/db_xref="GI:17384634"
/translation="MIFPLTDLIKPYCLMKLYEYKNKIDVLVLTEDIIPEDKLELIQ
SDDIVLARKKTCHELLURAKSKKLTESKRIARKIMARRERIRKNKIKLKEYN
KSIWKDIDQIELVLMQDMNKVHSSITNALHSLDTSRIINWDDLLNEVVRETLN
NIVAGAIKIRNPDKILDPGEANNIQLINDADAPLUNKIIEINEYMRITLPLEQINILL
NSFKENYLSIIQE"
7802..8455
/gene="escF"
7802..8455
/gene="escF"
/function="component of the type III secretion system"
/note="also known as sepi gene"
/codon_start=1

gene
CDS
/transl_table=11
/product="St04 protein"
/protein_id="CAC81842.1"
/db_xref="GI:17384630"
/translation="MISVVFIIINETGESLMINGLNNSASLVDAIRINSDFKKQ
WDMSCAKLLKLVLSFGWNETYRSRQTELLTVLEPVSAPNELGIYANFADG
SSLRVSNSELVAEINTPDNEKILMLLESNEQRLLQSLPINLHMPYIQVHRALSK
MDLTDHSMHNLSTFKLSATLPHINTQTDPLSGPTPFSSMFMDFRGNGAKLSLN
GVDPDAQKFRDAGLKDTHSSILARNVINGISRHHAKQIAPRESSGSKQKAEVVE
FLCHPEAATAICSAFYOSFNPALMLTHTRIQSAREYNVERSLEDVFNACINISISQSP
DGSITHVASTGILLIMAPEDRNELGMLTNTTSYEVPOGVKCEIDEMVRTILOPRYGASE
TYLKN1"
5612..6001
/gene="ler"
5612..6001
/gene="ler"
/function="putative transcription activator"
/codon_start=1
/transl_table=11
/product="ler protein"
/protein_id="CAC81843.1"
/db_xref="GI:17384631"
/translation="MRRLFIMNMTSHTTSPYIQLIEQLVLOQAKRLRGEIIOV
IESIQKQIYYNITLQELGYTNVDDGLARNSSKGVYRNNEGQTSWGVGRPWIK
EALLNGMKKEDELVKDTEDEIIPLKN1"
6016..6234
/gene="st06"
6016..6234
/gene="st06"
/function="putative secreted protein"
/codon_start=1
/transl_table=11
/product="St06 protein"
/protein_id="CAC81844.1"
/db_xref="GI:17384632"
/translation="MITITELEDEIKKNEAANIFTEKINDKKNETHKMKHPLDKVT
YNKAKELLACDAAIRIIEIMLIRINN"
6558..7157
/gene="st07"
6558..7157
/gene="st07"
/function="unknown"
/codon_start=1
/transl_table=11
/product="St07 protein"
/protein_id="CAC81845.1"
/db_xref="GI:17384633"
/translation="MTILNKIDYNYKLIYPIIMVHDEWLGLTGVNQVSRHILRES
SSTRQNLKILRQETQDKISGVELSDINKEGFLYQSIGKIRLLASSALFEIQPDYI
FSRLYELIIRIIGYONVKQLSFYWGQCKPEYGEERFCSELIKYAGNLEWLFSDN
PLMTIVKLLPKSGBIKPTINDLFLNRLNKLILPYETL"
7102..7797
/gene="st08"
7102..7797
/gene="st08"
/function="unknown"
/codon_start=1
/transl_table=11
/product="St08 protein"
/protein_id="CAC81846.1"
/db_xref="GI:17384634"
/translation="MIFPLTDLIKPYCLMKLYEYKNKIDVLVLTEDIIPEDKLELIQ
SDDIVLARKKTCHELLURAKSKKLTESKRIARKIMARRERIRKNKIKLKEYN
KSIWKDIDQIELVLMQDMNKVHSSITNALHSLDTSRIINWDDLLNEVVRETLN
NIVAGAIKIRNPDKILDPGEANNIQLINDADAPLUNKIIEINEYMRITLPLEQINILL
NSFKENYLSIIQE"
7802..8455
/gene="escF"
7802..8455
/gene="escF"
/function="component of the type III secretion system"
/note="also known as sepi gene"
/codon_start=1

gene
CDS
/transl_table=11
/product="St04 protein"
/protein_id="CAC81842.1"
/db_xref="GI:17384630"
/translation="MISVVFIIINETGESLMINGLNNSASLVDAIRINSDFKKQ
WDMSCAKLLKLVLSFGWNETYRSRQTELLTVLEPVSAPNELGIYANFADG
SSLRVSNSELVAEINTPDNEKILMLLESNEQRLLQSLPINLHMPYIQVHRALSK
MDLTDHSMHNLSTFKLSATLPHINTQTDPLSGPTPFSSMFMDFRGNGAKLSLN
GVDPDAQKFRDAGLKDTHSSILARNVINGISRHHAKQIAPRESSGSKQKAEVVE
FLCHPEAATAICSAFYOSFNPALMLTHTRIQSAREYNVERSLEDVFNACINISISQSP
DGSITHVASTGILLIMAPEDRNELGMLTNTTSYEVPOGVKCEIDEMVRTILOPRYGASE
TYLKN1"
5612..6001
/gene="ler"
5612..6001
/gene="ler"
/function="putative transcription activator"
/codon_start=1
/transl_table=11
/product="ler protein"
/protein_id="CAC81843.1"
/db_xref="GI:17384631"
/translation="MRRLFIMNMTSHTTSPYIQLIEQLVLOQAKRLRGEIIOV
IESIQKQIYYNITLQELGYTNVDDGLARNSSKGVYRNNEGQTSWGVGRPWIK
EALLNGMKKEDELVKDTEDEIIPLKN1"
6016..6234
/gene="st06"
6016..6234
/gene="st06"
/function="putative secreted protein"
/codon_start=1
/transl_table=11
/product="St06 protein"
/protein_id="CAC81844.1"
/db_xref="GI:17384632"
/translation="MITITELEDEIKKNEAANIFTEKINDKKNETHKMKHPLDKVT
YNKAKELLACDAAIRIIEIMLIRINN"
6558..7157
/gene="st07"
6558..7157
/gene="st07"
/function="unknown"
/codon_start=1
/transl_table=11
/product="St07 protein"
/protein_id="CAC81845.1"
/db_xref="GI:17384633"
/translation="MTILNKIDYNYKLIYPIIMVHDEWLGLTGVNQVSRHILRES
SSTRQNLKILRQETQDKISGVELSDINKEGFLYQSIGKIRLLASSALFEIQPDYI
FSRLYELIIRIIGYONVKQLSFYWGQCKPEYGEERFCSELIKYAGNLEWLFSDN
PLMTIVKLLPKSGBIKPTINDLFLNRLNKLILPYETL"
7102..7797
/gene="st08"
7102..7797
/gene="st08"
/function="unknown"
/codon_start=1
/transl_table=11
/product="St08 protein"
/protein_id="CAC81846.1"
/db_xref="GI:17384634"
/translation="MIFPLTDLIKPYCLMKLYEYKNKIDVLVLTEDIIPEDKLELIQ
SDDIVLARKKTCHELLURAKSKKLTESKRIARKIMARRERIRKNKIKLKEYN
KSIWKDIDQIELVLMQDMNKVHSSITNALHSLDTSRIINWDDLLNEVVRETLN
NIVAGAIKIRNPDKILDPGEANNIQLINDADAPLUNKIIEINEYMRITLPLEQINILL
NSFKENYLSIIQE"
7802..8455
/gene="escF"
7802..8455
/gene="escF"
/function="component of the type III secretion system"
/note="also known as sepi gene"
/codon_start=1

```

/transl_table=11
/product="EsCR protein"
/protein_id="CAC81847.1"
/db_xref="GI:17384635"
/translation="MSQLMTIGSQPIFLIIVFLLSLRPFVWVIGTSFLKISIVLGIL
KNAIGIQVPPNMAITSVSLITMTFIMSPIILQINDNISOEPINYTDSDFOKRVDEKI
LSPYSGFLEKNEKVEPERAAQKLGNEILKKDSIFILLPAFTMGOLEAAFKIG
FDLYLFFIAIDLIIISNLLALGHMHVSPVTISIPFKILLFILVGGWQKLFELLVVN"
8455. .8724
gene

Query Match      14.6%; Score 113.8; DB 1; Length 60073;
Best Local Similarity 52.7%; Pred. No. 2e-17;
Matches 247; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 208 tggtaggggttagtaccggagaggtgattattgggttttttaattgggttttggcgcg 267
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8925 TTGACAGGCATAGTGCTAAAGAAATAAGTAGGCTTTTCATTGGGTTATCATTTACT 8984

QY 268 gtcccttttggcgcggtgatatggcggtttctctgcttatactttacgtggcgcgaca 327
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8985 ATTCTTTTGGGCAATAGATGGCGCTGGGCAGATTATTGATACATTTAAGAGGTTCAACA 9044

QY 328 atgggttacgatatctcaattctacaatagaagctgaacccctcacttttggcttgctttc 387
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9045 ATATCTTCAATTTTAAACCGCTATTAAGTGATTCATCTCTATCAGCTGGTGTTATTTTA 9104

QY 388 agccagtttttgggtgttattcttcttataagcgcgcgcatggagtttatataaaacatt 447
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9105 TACCAATTTATCAGCTGTGATCTTTGTTATTTCATGTGGGATACAAAGCGTTCGGATAAG 9164

QY 448 ctgtatgagtcatacnaattttaccacccagggcggtactttattatttgaccggcaattt 507
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9165 TTGTATTTATCCTACGAGATATTACCATTACAAGCCGATATTGCATTCATTCGCTGCTTTA 9224

QY 508 ttaaaatataccaggcagagtgaggagcgttttatcaattatgtgtcagtttctctctt 567
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9225 GTAGATTTTATTTCGCTATCGGATTCATTTATTAACCTGATGTTATCATTTTCAATT 9284

QY 568 cctgccataatagtatggttagtcgcatctggcttttaggtcttttaaatcggtcggca 627
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9285 CCCATGATTATCGGTATCTCTTATGTGATATGGGTTTGGGTTCTCTCAATAAAACAGCG 9344

QY 628 caacaattgaatgtgttttcttctcctcaatgccgctcaaaagtatatgg 676
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9345 CCTCAGCTAAACGPATTTCACATTATCACTGCCGGTAAAGATTGTATAG 9393
```

Search completed: August 2, 2002, 02:49:38
Job time: 7953 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2002, 00:47:00 ; Search time 236.42 Seconds
(Without alignments)
5657.205 Million cell updates/sec

Title: US-09-809-524B-1
Perfect score: 779
Sequence: 1 atgcacacacaggaataatga.....gactgggttcacatctgtatg 779

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3473872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	779	100.0	779	22	Salmonella dublin
2	763	97.9	779	22	Salmonella typhimu
3	750.6	96.4	13417	17	Partial sequence o
C	4	207.4	26.6	297	Virulence factor w
5	60.2	7.7	801	21	Bordetella pertuss
6	60.2	7.7	35026	21	Bordetella pertuss
C	7	52.2	6.7	184	Virulence factor w
8	42.4	5.4	6866	24	Human immune syste
9	42.4	5.4	16914	24	Human gene regulat

10	41.8	5.4	6102	24	AAS61356	Human gene regulat
11	41.2	5.3	16842	22	AAS46411	Tumour suppressor
12	41.2	5.3	16842	24	AAS61335	Human gene regulat
13	40.6	5.2	17538	24	ABL33157	Human immune syste
14	40	5.1	26997	22	AAS46748	Tumour suppressor
15	39.8	5.1	5311	24	ABL33019	Human immune syste
16	39.6	5.1	707	22	ABL24056	Human breast cance
17	39.6	5.1	73334	22	ABL34125	Human immune syste
18	39.4	5.1	9741	20	ABL33323	Human immune syste
C	19	39	5059	20	AAX84332	Stealth virus nucl
20	39	5.0	17738	24	ABL33539	Human immune syste
21	39	5.0	19576	24	AAS61258	Human gene regulat
22	39	5.0	34769	22	AAS46775	Tumour suppressor
23	38.8	5.0	8136	24	ABL32355	Human immune syste
24	38.8	5.0	8143	24	ABL33897	Human immune syste
25	38.8	5.0	13814	24	ABL33193	Human immune syste
26	38.6	5.0	5296	24	ABL33285	Human immune syste
27	38.6	5.0	53585	20	AAX20251	Borrelia burgdorfe
28	38.4	4.9	6782	24	ABL32776	Human immune syste
29	38.4	4.9	8349	24	ABL33789	Human immune syste
30	38.2	4.9	6047	24	ABL33459	Human immune syste
31	38.2	4.9	6118	24	ABL33031	Human immune syste
32	38.2	4.9	9524	24	ABL32839	Human immune syste
33	38	4.9	6558	22	AAS46549	Tumour suppressor
34	38	4.9	15224	24	AAS61362	Human gene regulat
35	37.8	4.9	6189	22	AAS46599	Tumour suppressor
36	37.8	4.9	6189	24	ABL34592	Human metastasis a
37	37.8	4.9	13511	24	ABL32281	Human immune syste
38	37.6	4.8	5574	24	AAS63338	Chemically protrea
39	37.6	4.8	6381	24	ABL32967	Human immune syste
40	37.6	4.8	6381	24	ABL34319	Human metastasis a
41	37.6	4.8	7134	24	ABL32482	Human immune syste
42	37.4	4.8	11047	22	AAS45480	Chemically protrea
43	37.4	4.8	11047	22	ABL33985	Human immune syste
44	37.4	4.8	15649	22	AAS45396	Chemically protrea
45	37.4	4.8	20510	23	ABL19710	Drosophila melanog

ALIGNMENTS

RESULT 1
AAD20626
ID AAD20626 standard; DNA; 779 BP.
AC AAD20626;
XX
XX
DT 03-JAN-2002 (first entry)
XX
DE Salmonella dublin ssat gene.
XX
KW ssat; antibacterial; vaccine; non-rodent animal; immunity; mortality;
KW diarrhoea; milk production; bacterial infection; ds.
XX
OS Salmonella dublin.
XX
PN WO200170247-A2.
XX
PD 27-SEP-2001.
XX
XX
PF 13-MAR-2001; 2001WO-US08042.
XX
PR 17-MAR-2000; 2000US-190178P.
XX
XX (PHAA) PHARMACIA & UPJOHN.
XX
PI Lowery DE, Kennedy MJ;
XX
XX
DR WPI; 2001-639093/73.
XX
XX Vaccine composition useful for conferring protective immunity in a
PT non-rodent animal, comprises first attenuated, non-reverting mutant
PT Salmonella bacterium having two or more inactivated genes within SPI2

PT region -
 XX Claim 5; Page 48; 58pp; English.
 PS The invention relates to a vaccine composition comprising an
 XX immunologically protective amount of a first attenuated, non-reverting
 CC mutant *Salmonella* bacterium in which two or more genes within the SPI2
 CC region have been inactivated. The vaccine is useful for conferring
 CC protective immunity on a non-rodent animal, by administering the vaccine
 CC to the animal, such that an improvement in mortality, symptomatic
 CC diarrhoea, physical condition and milk production are provided. The
 CC vaccine is useful for reducing the amount or duration of bacterial
 CC shedding by about 10% or more during infection in a non-rodent animal
 CC e.g. cattle, sheep, goats, horses, pigs, poultry and other birds, cats,
 CC dogs and humans. The vaccine is useful for delivering a polypeptide
 CC antigen to an animal. The vaccine is also useful for providing benefit
 CC to veterinary and human community health. The vaccine is safe and
 CC efficacious live vaccine, which need not be administered at a very large
 CC doses. The present sequence is *Salmonella* dublin ssar gene.
 XX Sequence 779 BP; 185 A; 143 C; 161 G; 290 T; 0 other;
 SQ

Query Match 100.0%; Score 779; DB 22; Length 779;
 Best Local Similarity 100.0%; Pred. No. 1.6e-206;
 Matches 779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcacacaggttaagtgcttattgcttgcattgctgtgctgttatttattgacattta 60
 DB 1 atggcacacaggttaagtgcttattgcttgcattgctgtgctgttatttattgacattta 60

QY 61 agcctttcttatttacttccctattaaaagtgccagtttagggccgctcttttaagt 120
 DB 61 agcctttcttatttacttccctattaaaagtgccagtttagggccgctcttttaagt 120

QY 121 aatggcgtgcttattgcttacccttcccatattaccattaccacagagaagatt 180
 DB 121 aatggcgtgcttattgcttacccttcccatattaccattaccacagagaagatt 180

QY 181 atgatcatattgttaagattacagttggttaggttagtcacagagaggtgattatt 240
 DB 181 atgatcatattgttaagattacagttggttaggttagtcacagagaggtgattatt 240

QY 241 ggttttttaattgggttttgcggcgttcccttttggcgttgatagcggggttt 300
 DB 241 ggttttttaattgggttttgcggcgttcccttttggcgttgatagcggggttt 300

QY 301 ctgcttgatcatttgcgtgcccgaacatgggtacgattacaaattctacaaagaagct 360
 DB 301 ctgcttgatcatttgcgtgcccgaacatgggtacgattacaaattctacaaagaagct 360

QY 361 gaaacctcaacttttggctgttttcagccagttttgtgtgttttttttttataagc 420
 DB 361 gaaacctcaacttttggctgttttcagccagttttgtgtgttttttttttataagc 420

QY 421 ggcgcgatggagtttatataaacattctgtatgagtcataatattaccaccagg 480
 DB 421 ggcgcgatggagtttatataaacattctgtatgagtcataatattaccaccagg 480

QY 481 ggtactttatttattgacccgcaatttttaaaatataatccaggcagagtgagaacctt 540
 DB 481 ggtactttatttattgacccgcaatttttaaaatataatccaggcagagtgagaacctt 540

QY 541 tatcaattatgtgcagtttctcttctctgcccataatagtattgagccgacttg 600
 DB 541 tatcaattatgtgcagtttctcttctctgcccataatagtattgagccgacttg 600

QY 601 gctttaggtcttttaaatcggtcgccacaacattgattgtttttctctcctaagccg 660
 DB 601 gctttaggtcttttaaatcggtcgccacaacattgattgtttttctctcctaagccg 660

QY 661 ctcaaaagtatatggtttctactgacgctcctgatctcattcccttatgctcttcaac 720
 DB 661 ctcaaaagtatatggtttctactgacgctcctgatctcattcccttatgctcttcaac 720

Db 661 ctcaaaagtatatggtttctactgacgctcctgatctcattcccttatgctcttcaac 720

QY 721 tatttggtgaaagcgaataattttatttattctaaaagactggtttccactgtatg 779
 DB 721 tatttggtgaaagcgaataattttatttattctaaaagactggtttccactgtatg 779

RESULT 2
 AAD20627
 ID AAD20627 standard; DNA; 779 BP.
 AC AAD20627;
 DT 03-JAN-2002 (first entry)
 DE *Salmonella typhimurium* ssar gene.
 KW ssar; antibacterial; vaccine; non-rodent animal; immunity; mortality;
 KW diarrhoea; milk production; bacterial infection; ds.
 OS *Salmonella typhimurium*.
 PN WO200170247-A2.
 PD 27-SEP-2001.
 PF 13-MAR-2001; 2001WO-US08042.
 PR 17-MAR-2000; 2000US-190178P.
 XX (PHAA) PHARMACIA & UPJOHN.
 PI Lowery DE, Kennedy MJ;
 XX WPI; 2001-639093/73.
 Vaccine composition useful for conferring protective immunity in a
 non-rodent animal, comprises first attenuated, non-reverting mutant
Salmonella bacterium having two or more inactivated genes within SPI2
 region -
 Claim 5; Page 49; 58pp; English.
 The invention relates to a vaccine composition comprising an
 immunologically protective amount of a first attenuated, non-reverting
 mutant *Salmonella* bacterium in which two or more genes within the SPI2
 region have been inactivated. The vaccine is useful for conferring
 protective immunity on a non-rodent animal, by administering the vaccine
 to the animal, such that an improvement in mortality, symptomatic
 diarrhoea, physical condition and milk production are provided. The
 vaccine is useful for reducing the amount or duration of bacterial
 shedding by about 10% or more during infection in a non-rodent animal
 e.g. cattle, sheep, goats, horses, pigs, poultry and other birds, cats,
 dogs and humans. The vaccine is useful for delivering a polypeptide
 antigen to an animal. The vaccine is also useful for providing benefit
 to veterinary and human community health. The vaccine is safe and
 efficacious live vaccine, which need not be administered at a very large
 doses. The present sequence is *Salmonella typhimurium* ssar gene.
 Sequence 779 BP; 187 A; 143 C; 161 G; 288 T; 0 other;
 SQ

Query Match 97.9%; Score 763; DB 22; Length 779;
 Best Local Similarity 98.7%; Pred. No. 4.6e-202;
 Matches 769; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 atggcacacaggttaagtgcttattgcttgcattgctgtgctgttatttattgacattta 60
 DB 1 atggcacacaggttaagtgcttattgcttgcattgctgtgctgttatttattgacattta 60

QY 61 agcctttcttatttacttccctattaaaagtgccagtttagggccgctcttttaagt 120
 DB 61 agcctttcttatttacttccctattaaaagtgccagtttagggccgctcttttaagt 120


```

ET CDS 11407
FT /*tag= }
FT /note= "Putative end site of yscU gene."
XX WO9617951-A2.
XX 13-JUN-1996.
XX
XX PF 11-DEC-1995; 95WO-GB02875.
XX PR 05-MAY-1995; 95GB-0009239.
XX PR 09-DEC-1994; 94GB-0024921.
XX PR 31-JAN-1995; 95GB-0001881.
XX PA (RPMS-) RPMS TECHNOLOGY LTD.
XX
XX PI Holden DW;
XX
XX DR WPI; 1996-287194/29.
XX DR P-PSDB; AAR97244.
XX
XX PT Identifying virulence genes in microorganisms - by introducing
XX PT mutants with insertion inactivated genes into environment and
XX PT retrieval and analysis of mutants
XX
XX PS Claim 46; Figure 11; 131pp; English.
XX
XX CC A method for identifying a microorganism having a reduced adaptation
XX CC to a particular environment comprising the steps of: (1) providing a
XX CC plurality of microorganisms each of which is independently mutated by
XX CC the insertional inactivation of a gene with a nucleic acid comprising
XX CC a unique marker sequence so that each mutant contains a different
XX CC marker sequence, or clones of the said microorganism; (2) providing
XX CC individually a stored sample of each mutant produced by step (1) and
XX CC providing individually stored nucleic acid comprising the unique
XX CC marker sequence from each individual mutant; (3) introducing a
XX CC plurality of mutants produced by step (1) into the said particular
XX CC environment and allowing those microorganisms which are able to do so
XX CC to grow in the said environment; (4) retrieving microorganisms from
XX CC the said environment or a selected part thereof and isolating the
XX CC nucleic acid from the retrieved microorganisms; (5) comparing any
XX CC unique marker sequences in the nucleic acid isolated in step (4) to the
XX CC (2); and (6) selecting an individual mutant stored as in step
XX CC of the marker sequences as isolated in step (4). The products and
XX CC methods can be used for identifying virulence genes in microorganisms.
XX CC The mutant microorganisms can be used in vaccines or to screen for
XX CC drugs which reduce virulence or compounds useful for preventing,
XX CC ameliorating or treating infections in animals or plants. This
XX CC sequence corresponds to the first half of the virulence gene
XX CC cluster (VGC) 2 of Salmonella typhimurium.
XX
XX SQ Sequence 13417 BP; 3580 A; 2816 C; 3301 G; 3718 T; 2 other;

Query Match
Best Local Similarity 96.4%; Score 750.6; DB 17; Length 13417;
Matches 767; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 atggcacacaggttaaatgagtgcttattgcattgcttggtggtttatttcgaccatta 60
DB 9573 atggcacacaggttaaatgagtgcttattgcattgcttggtggtttatttcgaccattg 9632
QY 61 agcctttctttacttacttccctattaaaagtgccagtttagggcgctcttttacgt 120
DB 9633 agcctttctttacttacttccctattaaaagtgccagtttagggcgccacttttacgt 9692
QY 121 aatggcgcttgcattgcattacccttcccatattaccattaccattaccgagaagatt 180
DB 9693 aatggcgcttgcattgcattacccttcccatattaccattaccattaccgagaagatt 9752
QY 181 atgatgcattatggaagattacagttggttagggtagtcaccgagaggtattatt 240

```

```

DB 9753 atgatgcattatggaagattacagttggttagggtagtcaccgagaggtattatt 9812
QY 241 ggtttttaaattggttttgcggcggttcccttttgcggcggttgcattgagcgaggtttt 300
DB 9813 ggttttcaaattggttttgcggcggttcccttttgcggcggttgcattgagcgaggtttt 9872
QY 301 ctgcttgatacttactgcgtggcgacaaatgggtacgataattcaattcacataagaact 360
DB 9873 ctgcttgatacttactgcgtggcgacaaatgggtacgataattcaattcacataagaact 9932
QY 361 gaaacctcaacttttggcttgccttttcagcgaatttttgcgttgcatttcttcttaagc 420
DB 9933 gaaacctcaacttttggcttgccttttcagcgaatttttgcgttgcatttcttcttaagc 9992
QY 421 ggcgcgatggaggtttatatataaacattctctgtatgagtcataataattaccaccagg 480
DB 9993 ggcgcgatggaggtttatatataaacattctctgtatgagtcataataattaccaccagg 10052
QY 481 cgtactttattattgacggcgcaatttttaaaatatataccagggcagagtgagaaactt 540
DB 10053 cgtactttattattgacggcgcaatttttaaaatatataccagggcagagtgagaaactt 10112
QY 541 tatcaattatgctcagttctctcttctcgcataataatgattgattgagcagatctg 600
DB 10113 tatcaattatgctcagttctctcttctcgcataataatgattgattgagcagatctg 10172
QY 601 gcttaggtcttttaaatcggctgcggcacaacaaattgaattgtttttcttctcaatgcg 660
DB 10173 gcttaggtcttttaaatcggctgcggcacaacaaattgaattgtttttcttctcaatgcg 10232
QY 661 ctcaaaagtatttggttctactacgctcctgcattccattcccttatgtcttctcaac 720
DB 10233 ctcaaaagtatttggttctactacgctcctgcattccattcccttatgtcttctcaac 10291
QY 721 tattggtgaaagcgataaaattttattatttataaaagactggtttccatctgtatg 779
DB 10292 tattggtgaaagcgataaaattttattatttataaaagactggtttccatctgtatg 10350

RESULT 4
AAT09201/G
ID AAT09201 standard; DNA; 297 BP.
XX
XX AC AAT09201;
XX
XX DT 06-JAN-1997 (first entry)
XX
XX DE Virulence factor with similarity to E.coli fliQ and invX genes.
XX
XX KW Mutant; adaptation; virulence factor; identification; screening;
XX KW vaccine; drugs; infection; treatment; ss.
XX
XX OS Salmonella typhimurium.
XX
XX PN WO9617951-A2.
XX
XX XX 13-JUN-1996.
XX
XX PF 11-DEC-1995; 95WO-GB02875.
XX
XX PR 05-MAY-1995; 95GB-0009239.
XX PR 09-DEC-1994; 94GB-0024921.
XX PR 31-JAN-1995; 95GB-0001881.
XX
XX PA (RPMS-) RPMS TECHNOLOGY LTD.
XX
XX PI Holden DW;
XX
XX DR WPI; 1996-287194/29.
XX
XX PT Identifying virulence genes in microorganisms - by introducing
XX PT mutants with insertion inactivated genes into environment and
XX PT retrieval and analysis of mutants

```

XX Claim 32: Figure 6; 131pp; English.

XX A method for identifying a microorganism having a reduced adaptation

CC to a particular environment comprising the steps of: (1) providing a

CC plurality of microorganisms each of which is independently mutated by

CC the insertional inactivation of a gene with a nucleic acid comprising

CC a unique marker sequence so that each mutant contains a different

CC marker sequence, or clones of the said microorganism; (2) providing

CC individually a stored sample of each mutant produced by step (1) and

CC providing individually stored nucleic acid comprising the unique

CC marker sequence from each individual mutant; (3) introducing a

CC plurality of mutants produced by step (1) into the said particular

CC environment and allowing those microorganisms which are able to do so

CC to grow in the said environment; (4) retrieving microorganisms from

CC the said environment or a selected part thereof and isolating the

CC nucleic acid from the retrieved microorganisms; (5) comparing any

CC marker sequences in the nucleic acid isolated in step (4) to the

CC unique marker sequence of each individual mutant stored as in step

CC (2); and (6) selecting an individual mutant which does not contain any

CC of the marker sequences as isolated in step (4). The products and

CC methods can be used for identifying virulence genes in microorganisms.

CC The mutant microorganisms can be used in vaccines or to screen for

CC drugs which reduce virulence or compounds useful for preventing,

CC ameliorating or treating infections in animals or plants. This

CC virulence factor sequence was designated p9B7_1_1.

XX Sequence 297 BP; 99 A; 67 C; 55 G; 76 T; 0 other;

Query Match 26.6%; Score 207.4; DB 17; Length 297;

Best Local Similarity 97.2%; Pred. No. 6.5e-48;

Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 atggcacaacagggaaatgagtgcttattgcatggtggtggttttttttaccgaccatta 60

Db 224 ATGGCACACAGGTAATGAGTGCTTATTGTCATTGGCTGGCTTTATTTCGACCATG 165

Qy 61 agcctttctttattactccocctattaaaaagtgccagtggttagggcgctcttttcagt 120

Db 164 AGCCTTTCTTTATTACTCCCTTATTTAAAAAGTGGCAGTTTAGGGCGGCATTTCAGT 105

Qy 121 aatggcgctgattatgcattacattccattccattaccattaccattaccagagaagatt 180

Db 104 AATGGCGTGCTATGTGCTACCTTTCCGATATACCAATCATTTACCACGAGAAGATT 45

Qy 181 atgatgcatttggttaagattacagttggttaggt 217

Db 44 ATGATGCATATTGGTAAAGATTACAGTTGGGTAGGTT 8

RESULT 5

ID AAA64865

XX AAA64865 standard; DNA: 801 BP.

XX AAA64865;

XX 02-FEB-2001 (first entry)

DE Bordetella pertussis class I gene bact coding sequence.

DE bact: bacterial infection; anti-bacterial; vaccine; whooping cough;

KW type III secretion system; virulence factor; pathogenicity island; ss.

OS Bordetella pertussis.

XX Key Location/Qualifiers

PH CDS 1..801

FT /*tag= a

FT /product= "Bact"

XX W0200037493-A2.

PD 29-JUN-2000.

XX 21-DEC-1999; 99WO-EP10297.

PR 21-DEC-1998; 98GB-0028217.

XX (ULBR) UNIV LIBRE BRUXELLES.

XX Bollen A, Fauconnier A, Godfroid E;

PI WPI: 2000-452178/39.

DR P-PSDB: AAB14127.

XX Novel polypeptides derived from Bordetella pertussis, useful for

PT treating and diagnosing Bordetella infection -

XX Example 2; Pages 115-116; 165pp; English.

XX Bordetella pertussis possesses a type III secretion system. Type III

CC secretion systems allow bacteria to target virulence factors directly at

CC host cells. The present sequence is the bscI coding sequence of B.

CC pertussis. The present sequence is a Class I type gene and encodes a

CC protein involved in the type III secretion system of B. pertussis i.e.

CC a Bordetella pathogenicity protein. The present sequence is located

CC within a pathogenicity island (see AAA64890). A pathogenicity island is

CC a compact, distinct genetic unit carrying virulence genes. The protein

CC encoded by the present sequence may be used to treat or diagnose B.

CC pertussis infection, e.g. as a vaccine. Whooping cough is a disease

CC caused by infection by B. pertussis.

XX Sequence 801 BP; 105 A; 244 C; 259 G; 193 T; 0 other;

Query Match 7.7%; Score 60.2; DB 21; Length 801;

Best Local Similarity 47.9%; Pred. No. 7e-07;

Matches 235; Conservative 0; Mismatches 253; Indels 3; Gaps 2;

Qy 210 gttagggttagtcaccggagagggtgattggttttttaattgggttttgcggcggt 269

Db 219 gctggcgctgctggcgaaggaggcgtggtggcgatgcttcctgggttggctggcctt 278

Qy 270 tcccttttggcgctgatatggcggtttctgctgatactttacgtggcgcaaat 329

Db 279 gccatttggatcttcggaggccatcggtctgctcatagacaaccaacggcgccagcct 338

Qy 330 ggggtacgatatctcaattctacaatagagctgaaacctcaacttttggctgttttcag 389

Db 339 gggcgctatctcaaccccgccacgggcaacgattcgtcgccatgggcatctctcaa 398

Qy 390 ccagtttttggcttatttcttttataagcgggcgcatggagtttataaaacattct 449

Db 399 tctgggttcattggttctctctgacggcgggcggttgggttctcgccacgagct 458

Qy 450 gtatga-gtcatataattattaccaccagggcgactttattattatttggcggaatttt 508

Db 459 gtatgacagcttcgggttgggaacatctggcggtggtggcggtccatgccacaggg 518

Qy 509 taaaatat--atccaggcgagagtggagaaacgctttatcaattatgtcagttttctct 566

Db 519 cgccgtcggtgctggtgaccagttcagtggttttggcgcggtgctcgtgctggcctc 578

Qy 567 tctgccaataatgattggtatttagcgagactggctttaggtcttttaaacggtcggc 626

Db 579 gccggcactggtggcattgctctctggcgagctggcggtggcctgacgacgcttcgc 638

Qy 637 acaacaattgaatgttttttctctcaatgccgtcaaaaagtatatattggttctactgac 686

Db 639 gccctaacctcagggtgttctctcctggctcgtgccggtaaaagagcgcgctgttct 698

Qy 687 gctcctgatct 697

Db 699 gctggtgctgt 709

```

RESULT 6
AAA64890
ID AAA64890 standard; DNA; 35026 BP.
XX AC AAA64890;
XX DT 02-FEB-2001 (first entry)
XX DE Bordetella pertussis pathogenicity island coding sequence.
XX KW Bacterial infection; anti-bacterial; vaccine; whooping cough;
XX KW type III secretion system; virulence factor; pathogenicity island; ss.
XX OS Bordetella pertussis.
XX FH
XX Key Location/Qualifiers
FT CDS 711..2024
FT FT /*tag= a
FT FT /product= "Protein # 1 (AAB14147)"
FT FT complement (2055..3590)
FT FT /*tag= b
FT FT /product= "Protein # 2 (AAB14148)"
FT FT 4220..4696
FT FT /*tag= c
FT FT /product= "Protein # 3 (AAB14149)"
FT FT complement (4998..5948)
FT FT /*tag= d
FT FT /product= "Protein # 4 (AAB14150)"
FT FT 6160..6747
FT FT /*tag= e
FT FT /product= "Orf1 (AAB14137)"
FT FT complement (6783..7049)
FT FT /*tag= f
FT FT /product= "BscF (AAB14116)"
FT FT complement (7039..7338)
FT FT /*tag= g
FT FT /product= "BscE (AAB14115)"
FT FT complement (7379..8659)
FT FT /*tag= h
FT FT /product= "BscD (AAB14114)"
FT FT complement (8656..10755)
FT FT /*tag= i
FT FT /product= "BcrD (AAB14111)"
FT FT complement (10752..11120)
FT FT /*tag= j
FT FT /product= "Orf2 (AAB14133)"
FT FT complement (11117..11527)
FT FT /*tag= k
FT FT /product= "Orf3 (AAB14134)"
FT FT complement (11532..11909)
FT FT /*tag= l
FT FT /product= "Orf4 (AAB14135)"
FT FT complement (11906..13003)
FT FT /*tag= m
FT FT /product= "BopN (AAB14131)"
FT FT 13002..13784
FT FT /*tag= n
FT FT /product= "Orf5 (AAB14136)"
FT FT 13806..14081
FT FT /*tag= o
FT FT /product= "Orf6 (AAB14137)"
FT FT 14097..14582
FT FT /*tag= p
FT FT /product= "BcrH (AAB14112)"
FT FT 14630..15571
FT FT /*tag= q
FT FT /product= "Orf7 (AAB14138)"
FT FT 15601..16803
FT FT /*tag= r
FT FT /product= "Orf8 (AAB14139)"
FT FT 16827..17288
FT FT /*tag= s

```

```

/ product= "Orf9 (AAB14140)"
17293..17814
/ tag= t
/ product= "Orf10 (AAB14141)"
17892..18218
/ tag= u
/ product= "BscI (AAB14117)"
18215..19039
/ tag= v
/ product= "BscJ (AAB14118)"
19032..19694
/ tag= w
/ product= "BscK (AAB14119)"
19664..20302
/ tag= x
/ product= "BscL (AAB14120)"
20307..21641
/ tag= y
/ product= "BscN (AAB14121)"
21641..22150
/ tag= z
/ product= "BscO (AAB14122)"
22147..22695
/ tag= aa
/ product= "BscP (AAB14123)"
22692..23771
/ tag= ab
/ product= "BscQ (AAB14124)"
23768..24439
/ tag= ac
/ product= "BscR (AAB14125)"
24445..24711
/ tag= ad
/ product= "BscS (AAB14126)"
24723..25523
/ tag= ae
/ product= "BscT (AAB14127)"
25520..26569
/ tag= af
/ product= "BscU (AAB14128)"
26566..26964
/ tag= ag
/ product= "BscV (AAB14129)"
26955..28757
/ tag= ah
/ product= "BscC (AAB14113)"
complement (28778..29380)
/ tag= ai
/ product= "BrpL (AAB14130)"
complement (29412..29591)
/ tag= aj
/ product= "Orf11 (AAB14142)"
complement (29555..30529)
/ tag= ak
/ product= "Orf12 (AAB14143)"
30631..31776
/ tag= al
/ product= "Orf13 (AAB14144)"
complement (31773..33005)
/ tag= am
/ product= "Orf14 (AAB14145)"
32370..33014
/ tag= an
/ product= "Orf15 (AAB14146)"
complement (33002..34852)
/ tag= ao
/ product= "Protein # 5 (AAB14151)"

```

WO200037493-A2.
 29-JUN-2000.
 21-DEC-1999; 99WO-EP10297.

PF 15-MAR-2001; 2001WO-EP02955.
 XX
 PR 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPiG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2001-602752/68.
 XX
 PT Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer -
 XX
 PS Claim 1; SEQ ID No 133; 27pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (Ss) and sequences complementary to (ss). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 16842 BP; 3319 A; 710 C; 4940 G; 7873 T; 0 other;

Query Match 5.3%; Score 41.2; DB 22; Length 16842;
 Best Local Similarity 47.3%; Pred. No. 0.38;
 Matches 124; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 233 tgattatgggtttttaaattggtttgtgcggcggttccttttggccgttgatgg 292
 ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 14389 tgattatgggtattatttcgggtttttttatagtgctgggattataggcgtgagttatcg 14448
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 293 cgggggtttctgtgatactttacgtgcgcgcacaaatgggtacgatattcaattctacaa 352
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 14449 cgtttggttttttagttgtttttatgttttatgtttgttggtttttgttaagaattaaaa 14508
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 353 tagaagcgtgaacctcaacttttgggtgttttcagccaggtttttgtgtttttct 412
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 14509 aagaaaaaataatggagatggaatttatattgtatattgtttatgtttatgttg 14568
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 413 ttataagcggcggtgagttattataaacatttcgtatgaatcatcaattttac 472
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 14569 tataatttagtggtataagtatattatatacgttggtgtgtaattatttttaatttag 14628
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 473 caccaggcgtactttattatt 494
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 14629 aacgttcggtatttttaaaagt 14650
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12

AAS61335
 ID AAS61335 standard; DNA; 16842 BP.
 XX
 AC AAS61335;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human gene regulation-associated gene oligonucleotide #230.
 XX
 KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preeclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiac; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200177375-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP03968.
 XX
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPiG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 PT WPI; 2002-017470/02.
 XX
 PT New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease -
 XX
 PS Disclosure; SEQ ID No 297; 26pp; English.
 XX
 CC The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 CC preeclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 16842 BP; 3319 A; 710 C; 4940 G; 7873 T; 0 other;

Query Match 5.3%; Score 41.2; DB 24; Length 16842;
 Best Local Similarity 47.3%; Pred. No. 0.38;
 Matches 124; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy	233	tgattattggttttttaattgggttttctgtgcggcggttcctcttttgggcggttgatatgg	292
Db	14389	tgattttaggtgatttatttccggttttttattagtgtcgggattataggcgtgagttatcg	14448
Qy	293	cgggggttttcgcttgatacttttaagtcggcgcaaatgggtacgatatccaattctacaa	352
Db	14449	cgtttgggtttttagtctgtttttatgtttttatgttttgcgttgggttttttcttaagaattaaaa	14508
Qy	353	tagaagctgaacacacacttttttggcttgcgttttttcagccaggtttttgtgttattttct	412
Db	14509	aageaaaaaaaattatggagatggaattatattgtataattatctggtttatagtg	14568
Qy	413	ttataaaggcgcgcatggagtttatataaacaattctgtatagtcataatacattttac	472
Db	14569	tataatttagtggttaagtatattatatacgggttggtgtaattatttttaatttttag	14628
Qy	473	caccaggcgctactttattatt	494
Db	14629	abcgttcggtatttttaaaagt	14650
RESULT 13			
ABL33157			
ID	ABL33157 standard; DNA; 17538 BP.		
XX	XX		
XX	ABL33157;		
XX	XX		
DT	26-MAR-2002 (first entry)		
XX	XX		
DE	Human immune system associated gene SEQ ID NO: 1130.		
XX	XX		
KW	Human; immune system disease; cytosine methylation; antiasthmatic;		
KW	antiarteriosclerotic; antihaemic; cytostatic; nootropic;		
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;		
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;		
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;		
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;		
KW	gene; ds.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PN	W0200200928-A2.		
XX	XX		
PD	03-JAN-2002.		
XX	XX		
PF	02-JUL-2001; 2001WO-EP07537.		
XX	XX		
PR	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043826.		
XX	XX		
PA	(EPIG-) EPIGENOMICS AG.		
XX	XX		
PI	Olek A, Piepenbrock C, Berlin K;		
XX	XX		
DR	WPI; 2002-130909/17.		
XX	XX		
PT	Nucleic acid comprising fragment of chemically modified gene, useful		
PT	for diagnosis and treatment of diseases associated with abnormal		
PT	cytosine methylation		
XX	XX		
XX	Claim 1; SEQ ID NO 1130; 32pp + Sequence Listing; German.		
PS	XX		
CC	The present invention provides a number of human immune system associat		
CC	genes which are modified by the methylation of cytosines. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention.		
XX	XX		
SQ	Sequence 17538 BP; 5142 A; 188 C; 3355 G; 8853 T; 0 other;		

Query Match	5.2%	Score 40.6;	DB 24;	Length 17538;
Best Local Similarity	45.5%;	Pred. No.	0.57;	
Matches 185;	Conservative 0;	Mismatches 219;	Indels 3;	Gaps 17;

Qy	111	tcctttacgtaaatggcggtcattatgcacttaccctctcccatattaccatattacccaattacccta	170
Db	4940	tattttaagttagtgtgaatgtatatattagatttttaataaaattatgatataattataataa	4999
Qy	171	gcagaagaattatgatgcataatggttaaagattacaagttagggttagtgtaoaccggaga	230
Db	5000	ggtgaaaaataattgttaatagggatttaaaaataagtttaatttaattatagagaggtta	5059
Qy	231	ggtgatattatggtttttaattatgggtttcttgccggcggtt---cccttttgggccgttga	287
Db	5060	gtcttttagtctttttaaattctgaataataatagtttttttggaataattcttttaaaattta	5119
Qy	288	tatggcggggtttctgcttgatacttttacgtggcgcgacaatgggtgaagattccaattc	347
Db	5120	tgttgggttttattttaataaaattttgataggatgttaagttaggaataagattttttt	5179
Qy	348	tacaatgaagactgaanaacctcaccttttttgggttgccttttccagccagtttttgtgttat	407
Db	5180	taaaaggttaagtctatagttgaaataattttatattttttataaaagctattttagtaagta	5239
Qy	408	ttctttataagcggcgcatggagtttatataaaacattctctgatgcatacataa	467
Db	5240	ttttatttttatgttgtgttgatttttataatakatattaatttttaaaaggttttatgata	5299
Qy	468	tttaccacccagggcgctacttttattatttggaccggcaattttttaaagt	514
Db	5300	ttaaaaaattatagatgaataatttttagagtggtgtaatttttagagt	5346

RESULT	14
AAS46748	
ID	AAS46748 standard; DNA; 26997 BP.
XX	
AC	AAS46748;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Tumour suppressor gene derived chemically modified sequence #472.
XX	
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW	cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW	cytosine methylation; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200168912-A2.
XX	
PD	20-SEP-2001.
XX	
PF	15-MAR-2001; 2001WO-EP02955.
XX	
PR	15-MAR-2000; 2000DE-1013847.
PR	06-APR-2000; 2000DE-1019058.
PR	07-APR-2000; 2000DE-1019173.
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2001-602752/68.
XX	
PT	Fragments of chemically modified genes associated with tumour suppressor
PT	genes and oncogenes, useful in designing primers and probes for
PT	analysing diseases associated with cytosine methylation state e.g.
PT	cancer -

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2002, 00:41:25 ; Search time 63.82 Seconds
(without alignments)
2998.253 Million cell updates/sec

Title: US-09-809-524B-1
Perfect score: 779
Sequence: 1 atggcacacaggtaaatga.....gactgggttccatctgtatg 779

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750.6	96.4	13417	2	US-08-637-759B-37
2	750.6	96.4	13417	3	US-08-871-355A-37
3	750.6	96.4	13417	4	US-09-201-945-37
C 4	207.4	26.6	297	2	US-08-637-759B-15
C 5	207.4	26.6	297	3	US-08-871-355A-15
C 6	207.4	26.6	297	4	US-09-201-945-15
C 7	52.2	6.7	184	2	US-08-637-759B-16
C 8	52.2	6.7	184	3	US-08-871-355A-16
C 9	52.2	6.7	184	4	US-09-201-945-16
10	47.2	6.1	7218	1	US-08-232-463-14
C 11	37	4.7	852	4	US-09-586-933-6
C 12	33.6	4.3	90050	4	US-09-245-041-5
C 13	33.4	4.3	8133	1	US-08-480-604A-5
C 14	33.4	4.3	8133	2	US-08-405-496A-5
C 15	33.4	4.3	8133	4	US-08-915-136-5
16	32.4	4.2	1603	4	US-09-009-443-11
17	32.2	4.1	2676	1	US-08-525-596B-11
18	32.2	4.1	2676	3	US-09-177-860A-11
19	32.2	4.1	2676	3	US-08-891-789B-5
20	31.2	4.0	3528	4	US-08-984-320-2
21	31.2	4.0	3528	4	US-08-487-087A-2
22	31.2	4.0	4383	4	US-08-397-653B-2
23	31.2	4.0	4383	6	5175095-4
24	31.2	4.0	4383	6	517307-1
C 25	31.2	4.0	176373	3	US-09-128-155-17
C 26	30.8	4.0	485	3	US-09-284-783-17
C 27	30.8	4.0	569	4	US-08-936-165A-181

28	30.8	4.0	4673	1	US-07-638-431-1	Sequence 1, Appl
29	30.8	4.0	4673	5	PCT-US92-00018-1	Sequence 1, Appl
30	30.6	3.9	1620	4	US-08-858-207A-56	Sequence 56, Appl
31	30.6	3.9	2800	2	US-08-874-138-1	Sequence 1, Appl
32	30.6	3.9	2800	2	US-08-874-138-5	Sequence 5, Appl
33	30.6	3.9	2800	4	US-08-879-941-1	Sequence 1, Appl
34	30.6	3.9	2800	4	US-08-879-941-3	Sequence 3, Appl
35	30.6	3.9	2800	4	US-09-747-116-1	Sequence 1, Appl
36	30.6	3.9	2800	4	US-09-747-116-3	Sequence 3, Appl
C 37	30.4	3.9	7653	3	US-08-394-189B-1	Sequence 1, Appl
C 38	30.4	3.9	7653	3	US-08-258-287B-1	Sequence 1, Appl
C 39	30.4	3.9	7653	3	US-08-368-704C-1	Sequence 1, Appl
C 40	30.4	3.9	7653	5	PCT-US93-05701-18	Sequence 18, Appl
C 41	30.4	3.9	7653	5	PCT-US93-05703-1	Sequence 1, Appl
42	30.2	3.9	923	1	US-08-376-296-9	Sequence 9, Appl
43	30.2	3.9	923	3	US-08-910-691-10	Sequence 10, Appl
44	30.2	3.9	969	1	US-08-376-296-6	Sequence 6, Appl
45	30.2	3.9	969	3	US-08-910-691-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-637-759B-37
; Sequence 37, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: DNA sequence of VGC II from centre to left
; ORGANISM: hand end
; US-08-637-759B-37

Query Match	96.4%	Score 750.6	DB 2	Length 13417
Best Local Similarity	98.5%	Pred. No. 1.7e-208		
Matches 767	Conservative 1	Mismatches 10	Indels 1	Gaps 1
Qy 1	atggcacacagtgtaaatgaagtggttattgcattggcttggtctttttatttcgaccattta 60			
Db 9573	ATGGCACACAGGTAAATGAGTGGCTTATGCAATGGCTGGCTTTTATTTCGACCATTTG 9632			
Qy 61	agcctttctttattactctccctattaaagtggtgacgttttaggggcccgtctcttttaagt 120			
Db 9633	AGCCTTCTCTTATTACTTCTCCCTTATTAAGAGTGCACATTTAGGGCGCGCATTTTACGT 9692			
Qy 121	aatggcgtgcttatgtctacttaccctttcccatattaccatattaccagcagaagatt 180			
Db 9693	AATGGCGTGCCTATGTCTACTTACCTTCCGATATTAACCAATCATTTACAGCAGAAGATT 9752			
Qy 181	atgatgcataattggttaaagattacagttgggttaggggttagtcacgcgagaggtgattatt 240			
Db 9753	ATGATGCATATTGGTAAAGATTACAGTTGGTTAGGTTAGTCACTGGAGAGTGATTATT 9812			
Qy 241	ggtttttaattgggtttgtgcggcggttccctttttgggcgttgatagcgcggttt 300			
Db 9813	GGTTTTCAAATGGGTTTGTGCGCGGCTTCCCTTTTGGCGCTTGATATGGCGGGGTTT 9872			
Qy 301	ctgcttgatactttacgtggcgacgaatgggtacgatalattcaattctcaaatagaagct 360			
Db 9873	CTGCTTGATACCTTTACGTGGCGGACAAATGGGTACGATATTCAATCTCAATAAGAAGCT 9932			
Qy 361	gaacccctcaacttttggctgtgcttttcagccagtttttggcgcttgatagcggttt 420			
Db 9933	GAACCTCACATTTTGGCTTGTCTTTCAGCCAGTCTTGCTGTATTCTTTTATTAAGC 9992			
Qy 421	ggcgccatggagtttatataaacattctgtatgagtcataatcaatatattaccaccagg 480			
Db 9993	GGCGCATGGAGTTTATTAACATTTCTGTATGAGTCATATCAATATTACCACCAGGG 10052			
Qy 481	cgtaacttattattgaccggccaatttttaaaatataatccaggcagtgagagaagcgtt 540			
Db 10053	CGTACTTTATTATTATGACGACCAATTTTAAAAATATATCCAGGACAGTGGAGAACGCTT 10112			
Qy 541	tatacaattatgtctcagttctctcttcctcgccataatgtatggtattagccgatctg 600			
Db 10113	TATCAATATTATGATCAGCTTCTCTCTCCGCCATATATATGTATGTATGATGATGCGGATCTG 10172			
Qy 601	gcitttaggtcttttaaatcggtcgccacaacaattgaatgtgttttctctcctaatgcg 660			
Db 10173	GCITTAGGTCTTTAAATCGTGGCGACACAAATGAATGTGTTTTCTTCTCAATGCCG 10232			
Qy 661	ctcaaaagtatatgtgtctactgacgctcctgactcattcccttatgctcttcaatcac 720			
Db 10233	CTCAAAAGTATATGGTTCTACTGACG-YCCGATCTCAATTCCTTTATGCTCTTTCATCAC 10291			
Qy 721	tatttgggtgaagcgatataattttattatctaaagagctgggtttccactctglatg 779			
Db 10292	TATTTGGTTGAAGCGATAAAATTTATATTATCTAAAAAGACTGGGTTTCCATCTGTATG 10350			

RESULT 2

US-08-871-355A-37

: Sequence 37, Application US/08871355A

: Patent No. 6015669

: GENERAL INFORMATION:

: APPLICANT: David William Holden

: TITLE OF INVENTION: Identification of Genes

: NUMBER OF SEQUENCES: 501

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Patrea L. Pabst

: STREET: 2800 One Atlantic Center

: STREET: 1201 West Peachtree Street

: CITY: Atlanta

: STATE: Georgia

: COUNTRY: USA

```
Db 10053 CGTACTTTATTTGACGACCAATTTTAAATATATATCCAGGAGAGTGGAGACGCTT 10112
Qy 541 tatcaattatgtcaggtttctctctctccctgcccataaatatgtatggattagccgatctg 600
Db 10113 TATCAATATGATCAGCTTCTCTCTCCGCCATAATATGATGGTATTAGCCGATCTG 10172
Qy 601 gctttaggcttttaaatcggctcggcacaacaattgaatgtgttttctctctcaatgcg 660
Db 10173 GCYTTAGTCTTTTAAATCGGTGGGCAACAATTTGAATGTGTTTTCTTCTCAATGCCG 10232
Qy 661 ctcaaaagtatatgttctactgacgctcctgatctcattcccttatgctctctcatcac 720
Db 10233 CTCAAAAGTATATTTGGTCTTCTACTGACG-YCCTGATCTCATTTCCCTTATGCTTTCATCAC 10291
Qy 721 tatttgggtgaagcgataaattttatatctataaaagactgggtttccattctgtatg 779
Db 10292 TATTTGGTTGAAAGCGATAAATTTTATATTTATCTAAAGACTGGTTTCCATCTGTATG 10350
```

RESULT 3

```
US-09-201-945-37
; Sequence 37, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: DNA sequence of VGC II from centre to left
; ORGANISM: hand end
```

US-09-201-945-37

```
Query Match          96.4%; Score 750.6; DB 4; Length 13417;
Best Local Similarity 98.5%; Pred. No. 1.7e-208;
```

```
Matches 767; Conservative 1; Mismatches 10; Indels 1; Gaps 1;
Qy 1 atgcaacaacaggtaaatgagtgcttattgcattggtgctggtgctttatttcgaccatta 60
Db 9573 ATGCAACAACAGGTAATGAGTGGCTTATTGCAATGGCTGGCTTTTATTCGACCAATTG 9632
Qy 61 agcctttctttattacttccctattaaaaagtggcagtttagggggcgcgtcttttcagt 120
Db 9633 AGCCTTCTTTTATTACTTCCCTTATTAAAAAGTGGCAGTTTAGGGGCGGCACCTTTTAC 9692
Qy 121 aatggcgtcttatgtcacttacccttcccatattaccaatcattaccacagcagaagatt 180
Db 9693 AATGGCGTGTATGTCTACTTACCTTTCCGATATTACCAATCATTTACCACAGCAAGATT 9752
Qy 181 atgatcatattggttaaaagattacagttggttagggtttagtcacccggagagtgattatt 240
Db 9753 ATGATGCATATTGTTAAAGATTACAGTTGGTTAGGTTAGTCTACTGGAGAGGTGATTATT 9812
Qy 241 ggttttttaattgggttttctgcgcggttcccttttggcgcttggtgatatatggcgggttt 300
Db 9813 GGTTTTTCAAATTTGGGTTTGTGGCGCGGTTCCTTTTGGGCGCTTGATATGGCGGGGTTT 9872
Qy 301 ctgcttgatactttacgtggcgacaaatgggtacgatattcaattctacaaatagaagct 360
Db 9873 CTGCTTGATACTTTACGTGGCGCCACAAATGGGTACGATATTCAATCTTACAAATAGAAG 9932
Qy 361 gaaacctcaacttttggcttgcttttcagccagtttttctgctgtatttttcttataagc 420
Db 9933 GAAACCTCACCTTTTGGCTTCTTTTCAGCCAGTTCTTGTGTATTATTTCTTTATMAAG 9992
Qy 421 ggcggcatgagtttatataaaacattctgtatgagtcataatcaattttcccccagg 480
Db 9993 GCGGCGCATGGAGTTTATATTAAACATTCTGTATGAGTCATATCAATATTTTACCACGAG 10052
Qy 481 cgtactttatttgcacggcaatttttaaaatatataccaggcagagtgagaaacgctt 540
Db 10053 CGTACTTTTATTATTGACCAGCAATTTTAAATATATCCAGGAGAGTGGAGAAACGCTT 10112
Qy 541 tatcaattatgtcaggtttctctctctctccctgcccataaatatgtatggattagccgatctg 600
Db 10113 TATCAATATGATCAGCTTCTCTCTCCGCCATAATATGATGGTATTAGCCGATCTG 10172
Qy 601 gctttaggcttttaaatcggctcggcacaacaattgaatgtgttttctctcaatgcg 660
Db 10173 GCTTAGTCTTTTAAATCGGTGGGCGCACAAATTTGAATGTGTTTTCTTCTCAATGCCG 10232
Qy 661 ctcaaaagtatatgttctactgacgctcctgatctcattcccttatgctctctcatcac 720
Db 10233 CTCAAAAGTATATTTGGTCTTCTACTGACG-YCCTGATCTCATTTCCCTTATGCTTTCATCAC 10291
Qy 721 tatttgggtgaagcgataaattttatatctataaaagactgggtttccattctgtatg 779
Db 10292 TATTTGGTTGAAAGCGATAAATTTTATATTTATCTAAAGACTGGTTTCCATCTGTATG 10350
```

RESULT 4

```
US-08-637-759B-15/c
; Sequence 15, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Partial sequence of Salmonella typhimurium
; ORGANISM: virulence gene
; US-08-637-759B-15

```

```

Query Match      26.6%; Score 207.4; DB 2; Length 297;
Best Local Similarity 97.2%; Pred. No. 2.5e-51;
Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atggcacaacaggtaaatgagtgcttattgcattgctgtgcttatttcgacatta 60
   |||||||
Db 224 ATGCACACAGGTAATGAGTGGCTATTGCTATGCTGCTGCTTTATTTCGACATTG 165

QY 61 agccttcttattacttccctattataaaagtggcagtttagggcgctcttttaagt 120
   |||||||
Db 164 AGCTTTCTTTATTACTTCCCTATTATAAAAGTGGCAGTTTAGGGCGCGCATTTCAGT 105

QY 121 aatggcgctgttatgctacattcccttccattaccattaccattaccagagaagatt 180
   |||||||
Db 104 AATGGCGTGTATGTCACCTTACCTTTCCGATATTACCAATCATTTACCAGCAGAAGATT 45

QY 181 atgatgcatttggttaagattacagttggtagggt 217
   |||||||
Db 44 ATGATGCATATTGGTAAAGATTACAGTTGGGTAGGTT 8

```

```

RESULT 5
US-08-871-355A-15/c
; Sequence 15, Application US/08871355A
; Patent No. 6013569
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Partial sequence of Salmonella typhimurium
; ORGANISM: virulence gene
; US-08-871-355A-15

```

```

Query Match      26.6%; Score 207.4; DB 3; Length 297;
Best Local Similarity 97.2%; Pred. No. 2.5e-51;
Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atggcacaacaggtaaatgagtgcttattgcattgctgtgcttatttcgacatta 60
   |||||||
Db 224 ATGCACACAGGTAATGAGTGGCTATTGCTATGCTGCTGCTTTATTTCGACATTG 165

QY 61 agccttcttattacttccctattataaaagtggcagtttagggcgctcttttaagt 120
   |||||||
Db 164 AGCTTTCTTTATTACTTCCCTATTATAAAAGTGGCAGTTTAGGGCGCGCATTTCAGT 105

QY 121 aatggcgctgttatgctacattcccttccattaccattaccattaccagagaagatt 180
   |||||||
Db 104 AATGGCGTGTATGTCACCTTACCTTTCCGATATTACCAATCATTTACCAGCAGAAGATT 45

QY 181 atgatgcatttggttaagattacagttggtagggt 217
   |||||||
Db 44 ATGATGCATATTGGTAAAGATTACAGTTGGGTAGGTT 8

```

```

RESULT 6
US-09-201-945-15/c
; Sequence 15, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Partial sequence of Salmonella typhimurium
; ORGANISM: virulence gene
; US-09-201-945-15
```

```
Query Match 26.6%; Score 207.4; DB 4; Length 297;
Best Local Similarity 97.2%; Pred. No. 2.5e-51;
Matches 211; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atgcacacagatgaatgagtgcttattgacattgctgctgttatttcacacatta 60
|||||
DB 224 ATGGCACAACAGTAATAGTGGCTTATTGCAATGGCTGTGGCTTTATTCGACCAATTG 165
|||||

QY 61 agccttctttattacttcccttattaaagtgaggcttttagggcgctcttttacgt 120
|||||
DB 164 AGCCTTCTTTATTACTTCCCTTATTAAGTGGCAGTTTAGGGCGGCACATTTTACGT 105
|||||

QY 121 aatggcgctgttatgctacatttccatttccatttaccatttaccacggagagatt 180
|||||
DB 104 AATGGCGTGCTTATGTCACCTACCTTTCCGATATTACCAATCATTTTACCAGCAGAGATT 45
|||||

QY 181 atgatgcattgtgtaagattacagttgtagggt 217
|||||
DB 44 ATGATGCATATTGGTAAAGATTACATTGGGTAGGTT 8
|||||
```

```
RESULT 7
US-08-637-759b-16/c
; Sequence 16, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Partial sequence of Salmonella typhimurium
; ORGANISM: virulence gene
; US-08-637-759B-16

Query Match 6.7%; Score 52.2; DB 2; Length 184;
Best Local Similarity 67.1%; Pred. No. 2.6e-06;
Matches 104; Conservative 0; Mismatches 48; Indels 3; Gaps 2;

QY 46 ttatttcgaccattaaagccttcttttatttatttccctta--ttaaaaagtgagcagtttag 103
|||||
DB 175 TTAAGCGGGCCATTGACGCCCGCTGATACTGCGCTTAAGGAGAATGGGAGATTAAAG 116
|||||

QY 104 gggcgctcttttactgtaagtgcttattgctacatttaccatttcccaataattaccaatca 163
|||||
DB 115 GGGCGGGCTTTTACGTGATGACATACATATTGCACAAGCCTGGCCGATGTGTACCAATGA 56
|||||

QY 164 ttacca-gcagaagattatgatgcattatggtaa 197
|||||
DB 55 TTGCCACCCAGAGAAGATGATTGTTTCATAGTGGTTA 21
|||||

RESULT 8
US-08-871-355A-16/c
; Sequence 16, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871.355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

FILING DATE: 08/637,759
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/637,759
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.

TELECOMMUNICATION INFORMATION:


```

; TELEPHONE: (703)836-9300
; FAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgt-F1s
US-08-232-463-14

Query Match      6.1%; Score 47.2; DB 1; Length 7218;
Best Local Similarity 6.5%; Pred. No. 0.00031;
Matches 28; Conservative 216; Mismatches 184; Indels 0; Gaps 0;

QY 350 caatagaactgaacctcaatttttggctgcttttcagccagtttttggtgttatc 409
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1051 CGAGGAGCTCGATGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1110

QY 410 tctttaagcgcgcatggagtattataaaacattctgatgatcaaatatt 469
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1111 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1170

QY 470 tacaccaggcgctactttattatttagccggcaatttttaaaaatatccagcgag 529
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1171 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1230

QY 530 ggaagacgtttatcaattatgctgcagttctctcttcctgcgcataatgctatg 589
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1231 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1290

QY 590 tagccgatcgcttagtgtcttttaaactcgctggccaacaattgaatgttttct 649
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1291 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1350

QY 650 tctcaatgcgctcaaaagatatgttctactgacgctcctgatctcattcccttatg 709
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1351 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1410

QY 710 ctcttcatacatttgggtgaagcgataaaattttattattataaaagactggttc 769
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1411 YYYYYYYYYYYYYYYYYYYYYYYGTACCAAAATTCTCTACTCTTAACACTTG 1470

QY 770 catctgta 777
    |||||
DB 1471 GATAGGTA 1478

RESULT 11
US-08-586-935-6/c
Sequence 6, Application US/09586935
Patent No. 6191267
GENERAL INFORMATION:
APPLICANT: KONG, HUIMIN
APPLICANT: HIGGINS, LAUREN S.
APPLICANT: DALTON, MICHAEL
APPLICANT: KUCERA, REBECCA B.
APPLICANT: SCHILDKRAUT, IRA
TITLE OF INVENTION: Cloning And Producing The N.BstNBI Nicking Endonuclease
FILE REFERENCE: NEB-178
CURRENT APPLICATION NUMBER: US/09/586,935
CURRENT FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 6
LENGTH: 852
TYPE: DNA
ORGANISM: Pseudomonas lemoignei
FEATURE:
NAME/KEY: CDS

; LOCATION: {1}...(849)
US-09-586-935-6

Query Match      4.7%; Score 37; DB 4; Length 852;
Best Local Similarity 53.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 70;

QY 613 ttaaatcggctcggaacaattgaatgtttttctctctcaatgccgctcaaaagata 672
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 297 TTCAGTAGGATTAAATTCCTTCTAGTTCGTATTCAACTCTATTCTTTTCATAAATA 238

QY 673 ttggttctactgacgctcctgatctccatcccttatgtctcttcacactatttgg 732
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 TTCTATCTCATCAAGCTCATGACGAATTTGAACAAGAATTATCTTTAACATCTCGATAAAA 178

QY 733 agcgataaaatttatatttatctcaaaaaa 761
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 ATCTATAAGTTTTTTTATTAAATGTCATTGA 149

RESULT 12
US-09-245-041-5/c
Sequence 5, Application US/09245041
Patent No. 6274339
GENERAL INFORMATION:
APPLICANT: Moore, K.
APPLICANT: Nagle, D.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
OF BODY WEIGHT DISORDERS INCLUDING OBESITY
FILE REFERENCE: 7853-136
CURRENT APPLICATION NUMBER: US/09/245,041
EARLIER FILING DATE: 1999-02-05
EARLIER FILING DATE: 1998-07-21
EARLIER FILING DATE: 60/093,630
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 131
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
TYPE: DNA
ORGANISM: Mus musculus
US-09-245-041-5

Query Match      4.3%; Score 33.6; DB 4; Length 90050;
Best Local Similarity 70.3%; Pred. No. 7.4; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 19;

QY 353 tagaagctgaacctcacctcttttggctgttcttcagccagtttttggtgttatct 412
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 30935 TAGATCAAAAACACTTTCTTTTGTGTTTTTTTAAAGCAATTTTATTAGTATTCT 30876

QY 413 ttat 416
    |||
DB 30875 TCAT 30872

RESULT 13
US-08-480-604A-5/c
Sequence 5, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
```

```

; Patent No. 5919665
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, JAMES A.
; TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
; TITLE OF INVENTION: NEUROTOXIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405.496A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 424
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPD-01308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8130
; OS-08-405-496A-5

```

```

US-08-405-496A--5

Query Match          4.3%; Score 33.4; DB 2; Length 8133;
Best Local Similarity 52.5%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 338 tattcaattctacaatagaagctgaagaccctcaccttttggcttgcttccagcagtttt 397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1903 TATTTAATTCTAAATAGATTCTCCATCAACACATAAAAGTAGCTTTTGCACATTTTCAT 1844
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 398 tgtgtgtattttctttatagaagcggcggcattgagttattataaacattctgtatagtt 457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1843 TCATATTTCGTTGTATATATATACTATTTTGTAGGATTTTGTAGAAATAAATTGCATGTTG 1784
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 458 catatcaaatatttaccacc 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1783 CTTCAATAACTTATATCATC 1765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 15
US-08-915-136-5/c

Sequence 5, Application US/08915136
Patent No. 6290960
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 8133 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..8130
US-08-915-136-5

Query Match 4.3%; Score 33.4; DB 4; Length 8133;
Best Local Similarity 52.5%; Pred. No. 3.3;
Matches 73; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 338 tattcaattctacaagctgaagcctcaacttttttggttgcttgccttttcagccagtttt 397
Db 1903 TATTTAATCTAAATAGATTCTCCATCATCACTTAAAGTAGCTTTTTCACCTTTCAT 1844

Qy 398 tgtgtgtattttttttataaagcggcgcatggtttatattataaacattctgtatgagt 457
Db 1843 TCATATTTTCGTTGTATATAATACTATTTTAGGATTTTAGAAAATAAATTGCATGTTG 1784
Qy 458 catatcaatattttaccacc 476
Db 1783 CTTCAATAACTTATATATCATC 1765

Search completed: August 2, 2002, 02:48:30
Job time: 7625 sec

THIS PAGE BLANK (USPTO)